MULTIPLE DISCIPLINARY MODELING AND REPRODUCTION OF PHYSICAL PHENOMENA WITH COMPUTER SIMULATIONS AT HIGH AND LOW ATOMIC DETAIL

by

Robert Curtis DeMille

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

December 2012

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Robert Curtis DeMille**

has been approved by the following supervisory committee members:

Thanh Truong	, Chair	**02OCT12** Date Approved
Dmitry Bedrov	, Member	**02OCT12** Date Approved
Michael Morse	, Member	**02OCT12** Date Approved
Thomas Richmond	, Member	**02OCT12** Date Approved
Ryan Steele	, Member	**02OCT12** Date Approved

and by **Henry White** , Chair of

the Department of **Chemistry**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

A theoretical study of a chemical system is focused on representing the system properly with a model and using it to accurately represent and predict physical and dynamical properties of interest. The trade off between accuracy of simulations using a theoretical model and its computational expense is an important consideration in choosing and implementing the model and accompanying force field. My research has sampled the two extremes of this balance. In developing the mW-Ion and mW/3SPN-DNA models, a coarse-grained technique was used to simplify the interactions and significantly increase the efficiency of the calculations with respect to atomistic simulations. These models have limited transferability to other studies due to their coarseness, but reproduce properties such as solvation structure and ion dynamics quite well, and this with the ability to extend the simulation studies to timescales intractable for their atomistic counterparts. In later work, while investigating potential improvements to solid polymer electrolytes used in lithium battery technologies, an atomistic model with a polarizable force field was used in order to correctly capture the mobility of lithium cations. This involved a considerably larger computational expense, but was necessary to retain fidelity to experimental data. The advantages and disadvantages of the two sides of this balance is explored here, with detailed examination of the models and force fields used, their applicability, and broader impact in the simulation and scientific community.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to thank my advisors for their help and support during my time employed for them as a graduate student. I am also grateful to my friends and family for their support and encouragement. There are countless others, including fellow group members, who have helped me along the way, such as in providing computer and technical support, whom I wish also to acknowledge. I am appreciative for the advice and mentorship of my committee members, and especially of Ted Eyring, who, along with other faculty members, helped to give me perspective and motivation for completing my degree.

CHAPTER 1

INTRODUCTION

Molecular Dynamics

Since the advent of the adding machine, the computer has been used in enterprising endeavors to perform calculations too tedious and time consuming for a human being to complete. Computer models allow data to be processed, futures to be prognosticated, and understanding of the world around us to be refined, quantified, and read-out on a screen. The usefulness of a computer simulation, however, is only as good as the ability that it has to accurately represent the systems it attempts to model. As a model gets more detailed, more factors can be analyzed and predicted by the model, albeit at the cost of added computational overhead. As computers get faster and faster, systems can be more easily modeled, yet computing power is often paced by the size and complexity of the systems of focus. Thus, a necessary balance between the accuracy and the computational efficiency of a model must be taken into account when studying a system of interest.

Molecular Dynamics (MD) is a tool used by scientists for the purpose of understanding physical phenomena of chemical systems. Issues of both spatial and temporal resolution, typically concomitantly restricted using experimental techniques, can be overcome using MD simulations. MD can be applicable in many fields of study; simulations can offer a glimpse of solvation dynamics at an electrode surface,

fluctuations of biological membranes, crystallization of clathrates, and the motion of particles inside a carbon nanotube, to name a few cross-disciplinary examples. The folding of the bovine pancreatic tripsin inhibitor in 1975,[1] the first of its kind, showed in particular how powerful and how boundless investigations with MD could be.

MD works by treating atoms as spheres that undergo collisions with each other at the atomic scale and is governed by the equations for classical motion developed by Newton. At any given point in time of a MD simulation, the positions and velocities of all particles are known exactly. As the simulation progresses at each discretized timestep the forces acting on each particle are calculated and act to change the particles' momenta. This causes the particles to move accordingly to new positions until the next timestep, at which point the forces are calculated again and the process repeats. The timestep, typically on the order of femtoseconds, is chosen so as to capture the slowest degree of freedom of the system of study. Based on the starting position of the atoms, a trajectory of the particles results through the simulation dependent on the force field used, which is a mathematical equation governing how the particles interact with each other through bonded and non-bonded interactions. There are many computer programs that act to solve the equations of motions – LAMMPS[2], AMBER[3], and LUCRETIUS[4] are a few that I used in my research – each of which have different strengths and weaknesses mostly contingent on the type and size of the system of study as well as the underlying computer architecture used. The most important detail of the simulation, where its applicability and efficiency comes into play, is the force field.

The force field of a MD simulation, described as the potential energy U between atoms or groups of atoms, has the form of Equation 1.1, where the bonded interactions are typically comprised of bonds, angles, and dihedrals.

$$U = \sum_{bonds} U_{bond} + \sum_{angles} U_{bend} + \sum_{dihedrals} U_{torsion} + \sum_{nonbonds(i,j)} (U_{LJ} + U_{Coul}) \tag{1.1}$$

In my research, the bonded terms were not the focus of active exploration, using instead published values for the forms of these potentials and their respective force constants. I was actively interested in simplifying and improving through parameterization the non-bonded interactions between particles. Typical forms of short-ranged Lennard-Jones (LJ) and long-ranged Coulombic (Coul) potentials are found in Equations 1.2 and 1.3, respectively.

$$U_{LJ}(r_{ij}) = \left(\frac{A_{ij}}{r_{ij}}\right)^{12} - \left(\frac{B_{ij}}{r_{ij}}\right)^{6} \tag{1.2}$$

$$U_{Coul}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{1.3}$$

In these equations, A_{ij} and B_{ij} are Lennard-Jones constants for the pair of atoms i and j, with r_{ij}, being the distance between them, q is the charge of a particular atom, and ε_0 is the dielectric constant of a vacuum. Being dependent on the inverse sixth power of the

distance between the particles, hence quite short-ranged in terms of how quickly it approaches a value of effectively zero for increasing r, the Lennard-Jones interaction, U_{LJ}, is relatively inexpensive to calculate. The Coulombic interaction, U_{Coul},, however, decays inversely proportional to r, meaning it is inconveniently long-ranged and therefore must be calculated with care during an MD simulation to retain the fidelity of the simulation and to optimize its performance.

To calculate the Coulombic interaction during a simulation, typically a cutoff distance is chosen thereby ignoring the contribution to the potential energy of the system for atoms separated by a distance further than the cutoff. This has been shown, however, to introduce errors to the simulation due to it being a crude approximation, especially for biological systems.[5] To handle these long-ranged interactions, several methodologies have been introduced to calculate them in as much an error-free way as possible. The reaction field method[6] is one such technique, which considers the motion of an atom relative to a "cavity" of particles surrounding it for determining its contribution to the electrostatic potential. Ewald summation[7] and its offshoots particle mesh Ewald[8] and particle-particle-particle mesh[9] are, however, the most common methods, which employ mathematical "tricks" to optimize separately real- and reciprocal-space calculations using fast Fourier transforms of the Coulombic interaction. For a more detailed description of the Ewald method see Refs. [6, 7, 10-13] This method relies on using periodic boundary conditions for the simulation, meaning an infinite number of periodic images of the original system geometry are replicated in three dimensions. Whenever an atom moves across a periodic image, it appears on the other side. If the system size is too small, this can introduce errors in a MD simulation, as a particle could potentially "feel" itself across

a periodic image. The proper choice of the cutoff of the long-range interactions as well as the parameters of the Ewald summation routine can minimize these artificial errors.

Without using optimized methodologies to calculate the electrostatic interactions between atoms, the complexity of the calculation at each timestep scales as $N!$, with N being the number of particles. This is evident from Equation 1.3, where an increase in the total number of particles to four from two would result in six terms in the summation versus just one. This, of course, becomes prohibitively expensive in terms of computational overhead for any complex system. Ewald summation allows just an N^2 dependence, where particle-particle-particle mesh Ewald scales as $N \log N$, seemingly the best scaling possible with current electrostatic methodologies.

These methods for simplifying the electrostatic calculations are only applicable in the condensed phase; an inherent assumption of these models is that the effect of charges far away are screened out by an effective field of charges closer to an atom of interest. This screening is representative of the physical picture of interactions especially in the aqueous phase, where the average charge distribution beyond a close proximity is essentially net-neutral. The Yukawa potential[14] has been developed to replace the long-ranged Coulombic interaction and leverages this simplification of the potential due to the screening of charges. A cartoon representation of this physical picture is shown in Figure 1.1, where a positively charged ion is being pushed and pulled by water molecules and the other species in solution, effectively rendering long-ranged interactions to have a net effect of zero. This same physical picture of long-ranged interactions being effectively screened and reduced to just short-ranged interactions is what was employed in

developing a coarse-grained water model that was used extensively in the bulk of my research.

Coarse-Graining

Coarse-grain models have gained significant use in MD simulations for simplifying the amount of computation necessary.[15-22] By analyzing the system of study and understanding which of its components are the most important for replicating its observed physical properties, a model can be "coarsened" to build-in its less important details. A real-world example of this is the weather forecast: following the detailed minute-to-minute update of the temperature shows fluctuations of how one would feel outside, but a coarser description of the weather – that it will be warm today – is already often enough information to know what kind of clothes to wear. An example in a chemical system is treating a methyl group as particle of its own instead of a carbon atom bonded to three hydrogen atoms, a common coarse-grained approach in the OPLS – united atom model.[23] MD is already coarse-grained with respect to Quantum Mechanics (QM); details of the cloud of electrons surrounding atoms and comprising the bonds, wonderfully detailed and assiduously calculated with QM, is simplified and built in to the parameters used in MD simulations. Leaving out these details makes the simulation faster yet inherently less accurate, since not all factors are taken into account. Nevertheless, coarse-graining can be done in a way to maximize its applicability while retaining the desired level of fidelity to the factors deemed most important to the scientific study in question. The monatomic model of water developed by Moore and Molinero[24] illustrates this quite nicely.

The mW Model of Water

A bulk system of water has very defined physical properties of density, melting point, boiling point, etc. Just replicating all of these with MD simulations has proven to be difficult, engendering multiple models and iterations of force field parameter improvements. The TIP4P-Ew[25] model of water has generally been accepted as one of the most reliable at ambient temperature. Typical models of water such as TIP4P-Ew are all-atomistic (AA), meaning all the atoms are treated independently, and use long-ranged Coulombic interactions. These and the geometry of the water molecule give rise to the characteristic tetrahedrally-arranged hydrogen bond network also seen in experimental analyses of bulk water. The question that Moore and Molinero[24] posed in creating their coarse-grained model of water was essentially this: can an AA model of water be represented instead by a single particle with preferential tetrahedral hydrogen bonding effected through short-ranged potentials. Their idea behind the monatomic water (mW) model was that the physical picture of screened charges in solution could be leveraged to make the interactions short-ranged since a single water molecule only really "feels" its immediate neighbors in solution.

The force field of the mW model of water uses two- and three-body potentials to regulate the interactions between the spherically shaped, single particle water molecules according to Stillinger-Weber[26] (sw) potential, Equation 1.4.

$$E = \sum_i \sum_{j>i} \phi_2(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{k>j} \phi_3(r_{ij}, r_{ik}, \theta_{ijk})$$

$$\phi_2(r_{ij}) = A_{ij}\varepsilon_{ij}\left[B_{ij}\left(\frac{\sigma_{ij}}{r_{ij}}\right)^p - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^q\right]\exp\left(\frac{\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right) \tag{1.4}$$

$$\phi_3(r_{ij}, r_{ik}, \theta_{ijk}) = \lambda_{ijk}\varepsilon_{ijk}\left[\cos\theta_{ijk} - \cos\theta_{0ijk}\right]^2 \exp\left(\frac{\gamma\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right)\exp\left(\frac{\gamma\sigma_{ik}}{r_{ik} - a_{ik}\sigma_{ik}}\right).$$

In this equation, there are many adjustable parameters, the most important of which are the energy of the pair (ε_{ij}) and triplet (ε_{ijk}), the particle size (σ), the current angle (θ_{ijk}) and equilibrium angle (θ_{0ijk}) of a triplet of atoms, the three-body scaling factor (λ_{ijk}), and the value of $r_c=a\sigma$, which is the cutoff distance of the potential. During the simulation calculations, any triplet of atoms is penalized when not arranged with an angle between them corresponding to θ_{0ijk}. This provides the driving force to adhere to the tetrahedral angle, 109.48°, and allows for the anisotropy of the hydrogen-bonding network to be maintained with these simple monatomic water particles.

In coarse-graining the water molecule, several advantages are realized in the mW model with respect to typical AA models used. The two-body potential of the mW force field is "softer" than the Lennard-Jones equation, meaning it changes less dramatically as the distance between the particles fluctuates. This allows for a longer timestep to be used for the simulation – 10 fs versus 1 fs for TIP4P and others. Also, the lack of hydrogen atoms and the short-ranged nature (r_c=4.7Å) of the mW potential allow the calculations at each timestep to be sped up significantly. The three-body part of the mW force field is more time consuming than typical two-body calculations, but again, with a short cutoff distance, this penalty can be minimized. Moore and Molinero report[24] a 180-fold performance gain compared to the most efficient of fully atomistic water models.

Of course, a more efficient model of water does not necessarily equate with a better model. Intense parameterization efforts of the mW model were focused on ensuring fidelity to the physical properties of bulk water. The experimental values for the melting temperature, vaporization enthalpy, and density of liquid water were targets used by Moore and Molinero to fit the adjustable parameters of the sw equation. The density maximum anomaly and enthalpy of melting, two important aspects of water's behavior, are interestingly also captured by the mW model. The authors point out that the mW model fits just as well or better the physical properties of water as each AA model, showing how even the atomistic models often are not in agreement. Indeed, the mW model has been quite well accepted in the field, spawning offshoots using it in the Molinero group[27-30] and other independent work.[31] In my research, I used the mW model, its force field, and the underlying physical picture of the short-ranged nature of solvation in extending the model's functionality and applicability. These became the mW-Ion and the mW/3SPN-DNA models, which will be described in Chapters 2 and 3 in further detail as well as the parameterization efforts I used to develop them.

Polarizable Models

In the research towards the end of my graduate career, I was able to use MD simulations to validate and predict properties of chemical systems instead of the focus being on developing and parameterizing new models and force fields. An interesting application of simulation lies in the development of new materials to be used in commercial products. Engineering better materials for battery and fuel cell technology, without expensively creating and testing them first in the laboratory environment, can be aided by broad searches combining existing components with proposed improvements by

leveraging the raw number crunching power of computers. This is similar to drug design, another common modern direction of using the predictive powers of simulation and modeling. MD studies have been shown to be valuable in improving the fundamental understanding of ionic liquids (IL), important constituents of lithium batteries, as well as predicting properties of their mixtures.[32-35]

For MD simulations to be used in engineering technological improvements they, of course, must be able to accurately predict the properties of the system. This, again, depends on the quality of the force field. The Atomistic Polarizable Potential for Liquids, Electrolytes, & Polymers[36] (APPLE&P) was developed in the Smith/Bedrov group in order to address insufficiencies in other force fields particularly regarding applications of ionic liquids and lithium-ion battery constituents. Previously, typical studies of ILs using atomistic nonpolarizable force fields was problematic, with predictions of transport properties deviating by up to an order of magnitude from the experimental data.[37] Many of these force fields could represent well thermodynamic and structural properties of IL or molten salt systems, but underestimated the diffusion coefficient of the ions. This is due to the polarization, absent from these force fields, that a real chemical system undergoes in the presence of highly delocalized charge, in the case of ILs, or a high charge to size ratio, in the case of the Li^+ cation.

The Li^+ cation is highly disruptive to electron distributions of chemical species that surround it. Due to its small size, the electric field around Li^+ results in polarization of closely approaching solvating molecules. Approximately one third of the interaction energy between Li^+ and molecules that solvate it is attributable to this polarization.[36] If polarization is not taken into account, there is no fluctuation in the binding energy of this

interaction and allows the Li^+ to be tightly solvated since it can approach so closely. Only through decreasing artificially the positive charge of Li^+ by 10-20% and therefore the strength of interaction with its solvent and counter-ions can accurate predictions of Li^+ mobility be obtained for a two-body nonpolarizable force field.[32, 38, 39] Including induced dipoles to the APPLE&P force field allows for changes in the local electric field to be captured. These fluctuations help jettison Li^+ from its solvating species, thereby increasing its mobility to levels expected by experimental studies.

The induced dipole is implemented in the APPLE&P force field by calculating another term of the nonbonded energy, the polarization energy U^{pol},

$$U^{pol}(\mathbf{r}) = -\frac{1}{2}\sum_i \vec{\mu}_i \bullet \vec{E}_i^0 \qquad (1.5)$$

where $\vec{\mu}_i = \alpha_i \vec{E}_i^{tot}$ is an induced dipole at the force center *i*, $\vec{E}_i^{tot}$ is the total electrostatic field at the atomic site *i* due to the other sites' permanent charges and induced dipoles, α_i is the isotropic atomic polarizability, and $\vec{E}_i^0$ is the electric field due to fixed charges only. At each timestep, the induced dipoles are calculated in a self-consistent manner and act as an additional force influencing changes to the acceleration of each atom in the simulation. This calculation makes the polarizable force field more computationally expensive, requiring more overhead during the simulation routines by about a factor of four. The parameters of the force field are rigorously determined using quantum chemistry calculations; in particular the value of α_i is fit to reproduce the polarization response of a molecule due to a test charge.

The APPLE&P force field has been parameterized for a wide variety of ionic liquids and Li-battery components. The inclusion of polarization effects has been a main key to its success in predicting thermodynamic and dynamic properties. A recent study has shown that the polarization is vitally necessary; Bedrov, et al. investigated multiple ILs with the polarization both turned on and off for a side-by-side comparison of the physical properties and structure of the systems.[37] This study concluded that a 20-30% higher enthalpy of vaporization and a two to four times slower diffusion resulted from neglecting the polarization, while keeping the polarization allowed diffusion coefficients in particular to closely match with experimental values.

The lithium ion battery is a particular system of interest for study with simulation in order to improve its stability, capacity, and safety. The use of solid polymer electrolytes (SPEs) is an active area of research for applications in Li-battery technologies and proposes to address these issues by providing a polymer matrix for Li^+ to be dissolved in. Poly(ethylene oxide) (PEO) is a common SPE, yet can suffer from low conductivity at ambient temperature. By gaining a better understanding about the mechanism of transport of Li^+ and counter-ions in a matrix of PEO, Borodin and Smith hypothesized in a 2006 report using the APPLE&P force field,[40] better architectures could be proposed to improve the PEO SPE systems. They found that Li^+ motion is slow for intrachain motion but faster when transported between chains. A follow-up study[41] using a comb-branched PEO-based architecture was an attempt to give the cation more ability to make these chain-to-chain movements. The Li^+ mobility was found, however, to be lower due to slow movement along the backbone, especially for individual Li^+ cations that were coordinated at the oxygen-rich junction points where the backbone split

off forming the side chain combs. In Chapter 4 is presented the work that I did in providing another improvement to this comb-branched architecture by incorporating non-coordinating, flexible spacers at the attachment points of the SPE.

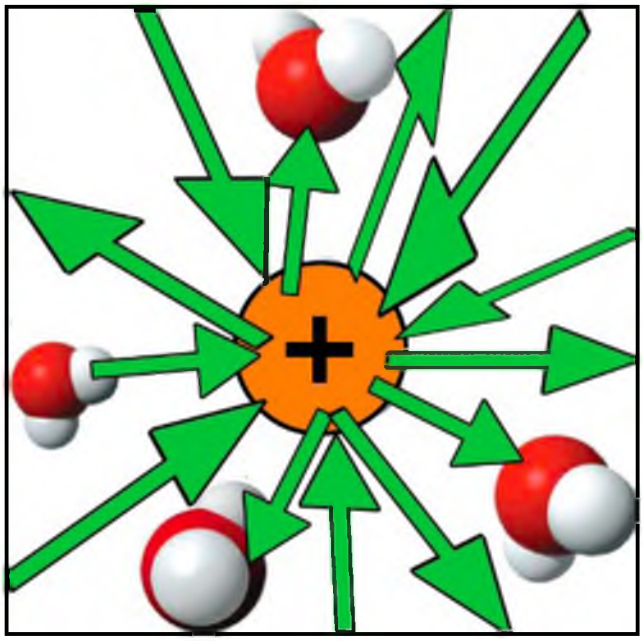


Figure 1.1. Cartoon of the idea of screening in condensed phase due to the effective cancellation of long-ranged attraction and repulsion due to charges felt more strongly by the cation at the center from the water molecules surrounding it closely.

CHAPTER 2

COARSE-GRAINED IONS WITHOUT CHARGES: REPRODUCING THE SOLVATION STRUCTURE OF NaCl IN WATER USING SHORT-RANGED POTENTIALS

This chapter was reproduced from the published paper with permission from DeMille, R. C. & Molinero, V. (2009) *The journal of chemical physics* **131**, 034107.

THE JOURNAL OF CHEMICAL PHYSICS **131**, 034107 (2009)

Coarse-grained ions without charges: Reproducing the solvation structure of NaCl in water using short-ranged potentials

Robert C. DeMille and Valeria Molinero[a)]
Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, USA

(Received 19 April 2009; accepted 16 June 2009; published online 17 July 2009)

A coarse-grained model of NaCl in water is presented where the ions are modeled without charge to avoid computationally challenging electrostatics. A monatomic model of water [V. Molinero and E. B. Moore, J. Phys. Chem. B **113**, 4008 (2009)] is used as the basis for this coarse-grain approach. The ability of Na^+ to disrupt the native tetrahedral arrangement of water molecules, and of Cl^- to integrate within this organization, is preserved in this mW-ion model through parametrization focused on water's solvation of these ions. This model successfully reproduces the structural effect of ions on water, referenced to observations from experiments and atomistic molecular dynamics simulations, while using extremely short-ranged potentials. Without Coulomb interactions the model replicates details of the ion-water structure such as distinguishing contact and solvent-separated ion pairs and the free energy barriers between them. The approach of mimicking ionic effects with short-ranged interactions results in performance gains of two orders of magnitude compared to Ewald methods. Explored over a broad range of salt concentration, the model reproduces the solvation structure and trends of diffusion relative to atomistic simulations and experimental results. The functional form of the mW-ion model can be parametrized to represent other electrolytes. With increased computational efficiency and reliable structural fidelity, this model promises to be an asset for accessing significantly longer simulation time scales with an explicit solvent in a coarse-grained system involving, for example, polyelectrolytes such as proteins, nucleic acids, and fuel-cell membranes. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3170982]

I. INTRODUCTION

The study of ionic species in aqueous medium has led to significant understanding of the physical processes at the atomic level. These results have far-reaching application in many fields, especially in understanding biological phenomena.[1–7] Molecular dynamics (MD) is an invaluable tool in gaining atomistic insight into these processes[8–10] and recent technological developments in high-performance computing have brought within reach the ability to bridge experiment and simulation.[11–15] Nevertheless, significant computational expense is seemingly unavoidable in evaluating the electrostatic effects present in ionic solutions.

Coulombic forces between particles decay as r^{-2}, with r being the distance between them, meaning they are notoriously long ranged relative to the dispersion forces. Electrostatics must be properly accounted for during MD simulations. The Ewald method[16] and its derivatives particle mesh Ewald (PME)[17] and particle-particle-particle mesh (PPPM)[18] have been developed and refined to contend with these long-ranged interactions. Other techniques such as the reaction field method[19] have been employed to simplify the treatment as well. For aqueous solutions, modeling the effect of long-ranged electrostatics has been further simplified by considering screened potentials that act over a shorter distance, thereby capturing the physical picture of an essentially net-neutral charge distribution beyond a close proximity. One such example is the screened Coulomb (Yukawa) potential.[20] Other modifications to shorten the length scale of the Coulomb potential include using a preaveraged Ewald potential[21] and "damp shifting."[22–24] Recently other studies have shown the ability to use a force-matching approach in fully atomistic simulations to replace the long-ranged electrostatics with potentials decaying in about 1 nm.[25,26] Each method has advantages and disadvantages in terms of computational cost and accuracy. The PPPM and PME methods remain likely as the most prevalent choices with a good trade-off between these two considerations[27] and widespread implementation.[28–31]

Coarse-grain (CG) approaches for simplifying atomistic simulations are employed with great success in the MD field.[32–53] Coarse graining involves careful analysis of the physics of the system of interest to extract effective interactions that can reproduce the behavior of the system for the properties of interest. This allows a simplified model to be developed that is computationally efficient yet still accurately portrays the behavior of the system at the level of detail required by the resolution of the coarse-grained model. One approach of CG studies, some involving explicit solvent, is to not coarsen ionic interactions, choosing instead to handle charged species with the full electrostatic treatment.[54–57] In the Mercedes-Benz (MB)-dipole model, for example, the two-dimensional MB water model was augmented with charges that produce a dipole, and this interacts with the charge of an ion through a long-ranged screened Coulomb potential.[58]

a)Author to whom correspondence should be addressed. Electronic mail: valeria.molinero@Utah.edu. Tel.: 1-801-585-9618.

Alternatively, electrostatics can be neglected entirely in a coarse-grained model, choosing to either implicitly define the interactions with the aqueous medium surrounding the system as a constant dielectric bath of water molecules or use a CG model of water without electrostatics. Some studies that have involved this latter approach include CG DNA, protein folding, and lipid membranes without charges.[32,59–61] In this paper we describe an approach for modeling ionic solutions where the charges themselves have been coarse grained; ions have been replaced by chargeless species, though interact with each other and with water to produce the solvation structure of charged species. This is based on, and is an extension of, the monatomic model of water (mW),[62] which has recently been developed in our group.

The coarse-grained model of water, mW, is based on the assumption that the interactions of a water molecule can be replaced by those of a single particle with anisotropic interactions. This coarsened molecule encourages a tetrahedral structure of the other molecules around it, akin to the native structure of liquid water effected by hydrogen bonding. The reader is referred to Ref. 62 for the full details of the mW model; only a brief highlight is given here. The mW model reproduces remarkably well the structure and thermodynamics of water in the condensed phase, including such details as the density maximum and phase transitions. Of particular note is the computational efficiency of the mW model, operating under 1% of the time required by atomistic models of water, yet showing comparable or better fidelity relative to experiment with regard to important physical properties such as the enthalpy of vaporization, density, and melting temperature. This performance gain results from the use of a longer time step, a reduction in the number of particles, and a shorter range of the intermolecular interactions compared to atomistic simulations.

In developing the mW model of water, a paradigm shift was made with regard to the range used for the interactions between molecules. Typical atomistic models of water, such as TIP4P-Ew, which we use as a reference in this study, involve assigning partial charges on or near the O and H atoms and using a competition of electrostatic attraction and short-ranged repulsion to produce hydrogen bonding. In other words, short-ranged ordering is effected through long-ranged interactions. Electrostatics are disregarded in the CG simulations presented here in favor of using short-ranged two-body and three-body (angular) nonbond interactions to influence the structure of water. In the mW model, short-ranged ordering is produced by short-ranged interactions.

Applying a CG approach to ions themselves in our simulations is a natural offshoot from observations gained during atomistic simulations and from literature reports. There have been many studies focused on the atomic structure of solvated ions,[63–66] some of which discuss, in particular, the so-called "structure-maker/breaker" effect.[67–71] These studies conclude, and with which atomistic simulations (detailed below) concur, that small cations such as Na^+ and K^+ disrupt the tetrahedral hydrogen-bond network of liquid water, while Cl^- is more readily integrated into this network. Inasmuch as the mW model of water is parametrized to encourage the tetrahedral structure of water using short-ranged and three-body potentials, mimicking the varying tetrahedrality of water around these ions meshes well with the tenets of the model schema. Again, this is performed while treating the ions as chargeless species using the structure of the solvated ions and water as the basis for parametrization of this mW-ion model.

Using conclusions garnered from atomistic simulations, we construct a model based on the coarse-grained mW and introduce the chargeless ions. Details of the model and the parametrization are found in Sec. II of this report. The methods used to calculate the structural, energetic, and dynamical properties from the simulations are also found in Sec. II. Our results are presented in Sec. III, as well as comparisons of our results to atomistic simulations and to other reports. These show the mW-ion model to accurately represent the atomic picture of solvated ions in bulk solutions even across a broad range of concentrations. Due to its very short-ranged nature, this model may find difficulty in describing systems in which long-ranged ordering is physically important, such as the asymmetrical interactions at an interface. The strength of the model, nevertheless, lies in its computational efficiency. The trade-offs between efficiency, fidelity, and transferability are discussed in Sec. IV, as well as an outlook of the mW-ion model.

II. METHODS

A. Atomistic reference simulations

To understand the effect of solvated ions on their surrounding water molecules, atomistic MD simulations were performed using the TIP4P-Ew (Ref. 72) model of water. We chose to use the TIP4P-Ew model due to its prevalence in the scientific literature and proven accuracy. There is a plentitude of ion parameters found in literature to describe the interaction of ions with atomistic models of water.[73–90] The Lennard-Jones (LJ) parameters used for the ions in our atomistic reference simulations were adapted from a recent report.[91] The ion parameters used in this work for the reference atomistic model are ε=0.1684 and 0.011 66 kcal/mol and σ=2.2589 and 5.1645 Å for Na^+ and Cl^-, respectively. The interaction energy between particles i and j is the sum of the LJ and Coulomb potentials, $E_{ij}=4\varepsilon_{ij}[(\sigma_{ij}/r_{ij})^{12}-(\sigma_{ij}/r_{ij})^{6}]+C\, q_i \cdot q_j/r_{ij}$, with r_{ij} being the distance between them and C an energy conversion constant, 332 kcal/mol. Geometric combination rules were employed for cross interactions, with $\sigma_{NaCl}=(\sigma_{Na}\sigma_{Cl})^{0.5}$ and $\varepsilon_{NaCl}=(\varepsilon_{Na}\varepsilon_{Cl})^{0.5}$. For the water-ion interaction, the same combination rules are used. 1000 water molecules plus ions comprised the simulation cell, and the LAMMPS[28] MD package was used to equilibrate the molecules and run the simulations with a 1 fs time step for 2.5 ns or longer under isobaric isothermal (npt) conditions at 298 K and 1 atm. The temperature was controlled with a Nosé–Hoover thermostat with a damping constant of 100 fs. Ions were added in separate simulations at the ratios (ion pair:water molecules) of 1:83, 1:40, 1:17, and 1:10, corresponding to the concentrations found in Ref. 68 for comparison. The PPPM method was used to calculate electrostatic interactions using 20 K-space vectors in each coordinate direction.

The mW-ion coarse-grained model presented in this work consists of two elements: the form of the interaction potential and its parameters. These are discussed in the next two subsections. The form of the potential constitutes the foundation of the model, and we expect it to be transferable to solutions of different electrolytes. The parameters presented here are specific to aqueous NaCl, more specifically to the reference atomistic model of aqueous NaCl used in this work. While the results of the coarse-grained parametrization depend on the choice of reference atomistic model, we measure the fidelity of the mW-ion against experimental data when these results are available.

B. Coarse-grained simulation models and force fields

The two main features of the potential energy forms used for the water-water, water-ion, and ion-ion interactions in the mW-ion model are (i) the use of three-body potentials to tune the number of first neighbors and hydrogen bonding (i.e., degree of tetrahedrality) of water-water and water-ion interactions and (ii) the use of extremely short-ranged potentials, which vanish at distances that range from 4.3 Å for water-water to 7 Å for ion-ion, to represent all the interactions in the model.

The main force field used for the coarse-grained simulations described here is the Stillinger–Weber[92] (SW) potential. Its form is given as

$$E = \sum_i \sum_{j>i} \phi_2(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{k>j} \phi_3(r_{ij}, r_{ik}, \theta_{ijk}),$$

$$\phi_2(r_{ij}) = A\varepsilon_{ij}\left[B\left(\frac{\sigma_{ij}}{r_{ij}}\right)^p - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^q\right]\exp\left(\frac{\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right),$$

$$\phi_3(r_{ij}, r_{ik}, \theta_{ijk}) = \lambda_{ijk}\varepsilon_{ijk}[\cos\theta_{ijk} - \cos\theta_{0ijk}]^2 \times \exp\left(\frac{\gamma\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right)\exp\left(\frac{\gamma\sigma_{ik}}{r_{ik} - a_{ik}\sigma_{ik}}\right), \quad (1)$$

Of note when examining this potential is the use of a three-body term, ϕ_3, in order to favor tetrahedral angles ($\cos\theta_{0ijk} = -1/3$) between triplets of particles (i, j, and k). The large number of parameters of this force field allows for a wide customization of the interaction behavior. Three parameters, in particular, are useful to exert this control. The energy scale ε influences the depth of the interaction, while σ determines the particle size. The parameter λ tunes the strength of the tetrahedral interactions by applying an energy penalty to configurations in which the three involved particles are within the cutoff distance $a\sigma$ and form an angle between them other than the specified tetrahedral angle $\theta_{0ijk} = 109.48°$. For the mW-ion model, the parameters A, B, p, q, a, θ_0, and γ in Eq. (1) remain fixed for all pairs and triplets of interactions with the values $A = 7.049\,556\,277$, $B = 0.602\,224\,558$, $p = 4$, $q = 0$, $a = 1.8$, $\theta_0 = 109.48°$, and $\gamma = 1.2$. With these parameters, the two-body potential has a minimum energy $-\varepsilon$ at a distance of 1.12σ. The forces calculated using the SW potential between pairs and triplets vanish at the cutoff distance of $r_c = a\sigma = 1.8\sigma$.

1. The mW model

The full details of the monatomic model of water are given in Ref. 62. It was parametrized to fit experimental vaporization enthalpy, melting temperature, and density of liquid water at room temperature. The model reproduces important features such as the radial and angular distribution functions of liquid water, enthalpy of melting of ice, and liquid anomalies (temperature of maximum density, increase in heat capacity upon cooling, and increase in diffusivity upon compression). The values of the parameters for the mW system using the SW potential are given in Table I. The angle $\theta_{0ijk} = 109.48°$ is used in this model to promote the tetrahedrality of the coarse-grain water molecules. The mW potential is softer than the typical LJ form and allows the use of a larger simulation time step of 10 fs. This, when combined with the short-ranged nature of the SW potential and the lack of hydrogens, allows a 180-fold performance gain compared to MD simulations with atomistic rigid water model.

2. Coarse-grain mW-ion model

The mW-ion model uses the SW force field [Eq. (1)] for the interaction of mW with the ions. During parametrization, described below, the degree of tetrahedrality of water surrounding either Na or Cl was adjusted through the parameter λ while using the same tetrahedral angle $\theta_{0ijk} = 109.48°$. Adjustment of λ was the primary method by which we were able to replicate the differences in solvation of Na and Cl ions. The full set of final parameters for the mW-ion model can be found in Table I. Using a range of time steps to find an optimum value required for this model, we performed several 5×10^6 step microcanonical simulations. Energy conservation ranged between 0.0066% for 2 fs time step and 0.18% for 10 fs. With energy conservation at 0.011%, we recommend a time step of 5 fs for this potential. This is half of the time step allowed by the mW water model; the reduction arises due to the fact that the interaction of the ions with water (ε) is larger than the water-water interaction and that the time step scales with the inverse of the interaction energy.

An additional force field, a shielded Coulomb (Yukawa) potential[20] is added to the interaction between ions of the same sign. The added Yukawa force field serves the purpose of supplying a repulsive force to Na–Na and Cl–Cl pairs. The resulting interaction is short ranged, which is justified by the physical picture of ions in aqueous medium. Although the Coulombic forces act over a long distance, they are shielded by the net effect of all of the surrounding charges (water's partial charges). The Yukawa potential has the form

$$E = Ar^{-1}\exp(-\kappa r), \quad r < r_c. \quad (2)$$

For the attraction of Na–Cl pairs, the SW potential remains to describe the interaction. This will be explained in further detail below in the section involving parametrization of the water-mediated ion-ion interaction.

C. Parametrization of the mW-ion model

Upon introducing coarse-grained ionic species into the mW model, parametrization of the structure of the ions and

TABLE I. Force field parameters of the mW-ion model.

Pair parameters for Eq. (1)[a]

i	j	ε_{ij} (kcal/mol)	σ_{ij} (Å)	a_{ij}
mW	mW	6.189	2.3925	1.8
mW	Na	16.00	1.85	1.8
mW	Cl	15.00	2.60	1.8
Na	Cl	10.2	2.00	1.8

Pair parameters for Eq. (2)

i	j	A (kcal/mol)	κ (Å^{-1})	r_c
Na	Na	1107	1.8	7.0
Cl	Cl	1107	1.8	7.0

Three-body parameters for Eq. (1)[b]

i	j	k	ε_{ijk}	λ_{ijk}	$\cos(\theta_{0ijk})$
mW	mW	mW	6.189	23.15	−1/3
mW	mW	Na	16	7	−1/3
mW	mW	Cl	15	16	−1/3
mW	Na	Na	16	7	−1/3
mW	Na	Cl	18.2	1	−1/3
mW	Cl	Cl	15	16	−1/3
Na	mW	mW	16	7	−1/3
Na	mW	Na	16	7	−1/3
Na	mW	Cl	18.2	1	0.746 06
Cl	mW	mW	15	16	−1/3
Cl	mW	Na	18.2	1	0.876 48
Cl	mW	Cl	15	16	−1/3

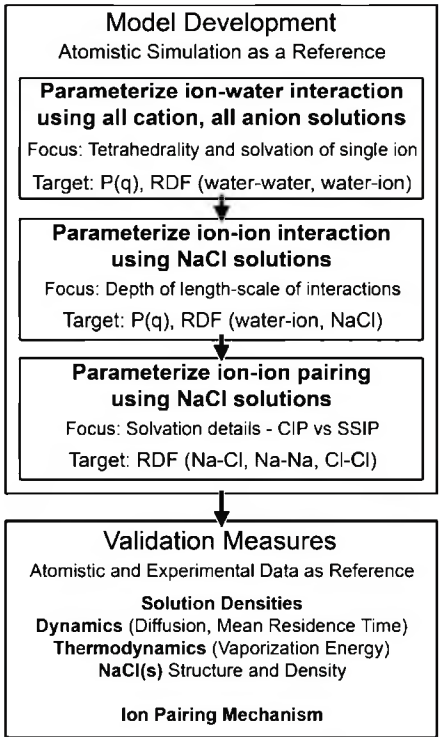
[a]Two-body parameters A=7.049 556 277, B=0.602 224 558, p=4, q=0, and γ=1.2 same for all ij pairs interacting through the SW potential [Eq. (1)].
[b]i is the center atom, with θ_0 forming the angle between legs r_{ij} and r_{ik} (ε_{ijk} are in units of kcal/mol). Three-body parameters for configurations ijk are the same as ikj. Terms not listed in the table (e.g., three-body terms for triplets of ions) do not contribute to the energy.

water was performed in order to reproduce that of the atomistic simulations. The simulation conditions used for the parametrization were the same as for the TIP4P-Ew simulations, involving 1000 water molecules and ions in the *npt* ensemble at 298 K and 1 atm. The development of the model was primarily focused on the structural changes brought about by having these "ions" present. A trial-and-error method was used primarily while developing the model, and we empirically determined how well the mW-ion results matched our reference data. Scheme 1 presents the flowchart of the parametrization procedure. Using dilute atomistic NaCl solutions, we first parametrized the cation-water and anion-water interactions. The focus here was to reproduce the solvation structure of water around the ions and to differentiate between them. In the second stage, we combined cations and anions together in the coarse-grained water. Here the length scales and depth of interaction between the ions both for attraction and repulsion were the focus of the parametrization efforts so as to reproduce the atomistic solvation structure of water around the ions. In the final step, the balance between ion-pairing and ionic solvation, witnessed in our atomistic simulations, was investigated and encoded into the mW-ion model in order to resolve both contact and solvent-separated ion pairs.

In studying the structure of these solutions, the radial distribution function (RDF), number of solvent shell neighbors, distribution of orientational order parameter $P(q)$, and angular distribution $P(\theta)$ were the bases for the model development (see Sec. II D, below). We developed the mW-ion model by comparing our results with these data (i.e., the training set) of our atomistic simulations. The solution density was not considered in the parametrization and was later used as a measure of the validity of the force field parameters used for the CG-ion model. The trends of features such as the self-diffusion coefficient and association/dissociation time were also used as a comparison to atomistic simulations and experiment to assess the ability of the model to capture the dynamical processes at play. The energy of vaporization was evaluated as a further measure of the ability of the coarse-grained model, trained exclusively to reproduce solvation structure, to capture the energetics of the system.

1. Ion-mW interaction

The SW potential was chosen for governing the interaction of water with the ions in order to leverage the ability of the three-body factor to fine tune the tetrahedral character of water around the ions. Of primary importance in developing



SCHEME 1. mW-ion model parametrization scheme.

the ion-mW interaction was studying the effects of changing three parameters of the SW potential: σ, the diameter of the particle; ε, the depth of the potential involving the interaction of an ion with mW; and λ, the degree of tetrahedrality of mW particles around an ion. This was done for both mW-cation and mW-anion dilute (1:83) solutions separately. Using for the coarse-grained model typical ionic radii utilized in atomistic simulations, the resulting RDFs showed maxima at larger distances than expected. With the atomistic simulation results as a reference, we optimized the ionic radii to their final values, found in Table I, so that the positions of the local minima and maxima of the RDFs were aligned closely with the atomistic data (Figs. 2–4 and 7, below).

Furthermore the combination of ε and λ was used to tune the strength and tetrahedrality of the ion-mW interaction. By increasing λ, a more tetrahedral arrangement is favored, showing more of a character like that of Cl^-. Conversely, decreasing λ allows for a lessened tetrahedral character like that of Na^+ to be observed for the CG-Na. Varying the interaction strength allowed the RDFs to more closely resemble the atomistic data, particularly regarding the agreement of the first shell solvation neighbors. Favoring the mW-ion interaction and disfavoring the ion-ion interaction through adjusting ε avoided clustering of ions. Since the conditions of the simulation were at high dilution, comparing an equally dilute atomistic solution is reasonable. To further clarify this comparison, when tuning the tetrahedral character of these solutions with only one ion type, configurations of the atomistic particles were not used as a reference if they included two or more ions in the first or second solvation shell of a water molecule. These final parameters we present in Table I represent our closest empirical match to the structural effects found in atomistic simulations.

2. Ion-ion interaction: ++ and − − repulsion

Reproducing the direct solvent-mediated electrostatic repulsion and attraction of the ions present in solution was found to be the most challenging interaction to mimic. Based on the LJ parameters used in this study for a TIP4P-Ew NaCl solution, the typical interaction depth of a contact ion pair (CIP) comprised of the LJ and Coulombic contributions is around 115 kcal/mol when the ions are at a the distance of the first peak of the RDF. Applying this interaction directly in the mW system would clearly dominate the energy landscape since the mW-mW interaction energy is less than 6 kcal/mol, as seen in Table I. In a similar fashion, using a long-ranged potential for the repulsion of the similarly "charged" ions would dominate the interactions, particularly at high concentration. This was, in fact, attempted with disastrous results, leading to a highly unphysical segregation of ions from the water. The Yukawa potential [Eq. (2)] was adopted to handle the ion-ion repulsion. Tuning the parameters of the Yukawa potential allowed it to be extremely short-ranged compared to the unshielded Coulombic interaction. For the mW-ion model, the values $A=1107$ kcal/mol and $\kappa=1.8$ Å^{-1} are used as the Yukawa parameters for interactions involving the pairs Na–Na and Cl–Cl with a cutoff of 7 Å. The range of the direct ion-ion interaction in the coarse-grained model is comparable to that between mW water particles, 4.3 Å.

3. Water-mediated and direct cation-anion interaction

The balance of cation-anion attraction and their solvation by water leads to two well characterized ion pair structures in aqueous solutions.[71,93,94] These are the CIP and the solvent-separated ion pair (SSIP), illustrated in Fig. 8(a). The electrostatic forces present between the partial charges on a water molecule and a sodium chloride pair give rise to the SSIP local minimum in the free energy landscape. This dipolar orientation is implicit to the TIP4P-Ew model of water; SSIPs occur naturally during the atomistic simulations. Excluding electrostatics from the coarse-grained ion model, however, leaves no inherent tendency to form SSIPs. Indeed, there is, in fact, no directional orientation for an mW water molecule besides favoring tetrahedral interactions with solvating neighbors. The differentiation of CIP and SSIP with a monatomic water proved to be one of the most challenging tasks during the parametrization. The key lie in the geometry of SSIP formation: we use explicit cation-water-anion three-body interactions to favor the geometry of triplet configurations seen in the atomistic simulations [e.g., Fig. 8(a)]. These

three-body interactions (entries "mW Na Cl," "Na mW Cl," and "Cl mW Na" from Table I) provide a balance to the direct attraction of Na and Cl governed by the two-body term (ϕ_2) of the coarse-grained potential.

Parametrizing this balance of attraction and repulsion of ion pairs was key to capturing the nuances of solvation structure. This will be further detailed below in the discussion regarding contact and solvent-separated ion pairs, highlighting the geometric considerations required to differentiate the two. While we used the Yukawa potential to model the repulsion of similarly charged ions, it was not necessary to describe the attraction of oppositely charged ions since the parametrization of the balance between CIPs and SSIPs already takes this attraction into account. With these refinements, all of the interactions experienced by the species present in solution are effectively zero at a distance around 6 Å.

D. Property calculations

Characterizing the ability of the CG-ion model to reproduce results matching our target atomistic reference was achieved through postprocessing of the configurations generated during the simulations. The methods used in this study to quantify the structural and dynamical properties of the model systems are presented here.

1. Structural

To examine the tetrahedral environment around the particles during the simulation, the orientational order parameter[95,96]

$$q_i = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\left(\cos\theta_{jk} + \frac{1}{3}\right)^2 \tag{3}$$

was used, where the four closest particles in the first shell comprise the set of j's and k's. These were summed for all particles i over the entire simulation and binned to give a distribution $P(q)$. For a high degree of tetrahedral order, a value close to 1 is expected, whereas a random distribution (as in an ideal gas) would correspond to $q=0$. Different distributions were obtained for cases in which ions are and are not present in the first shell.

The RDF is another important measure of the structure of the simulated solutions. Using the standard definition[97]

$$g_{ij}(r) = \frac{V}{N^2}\left\langle \sum_i \sum_{j\neq i} \delta(r - r_{ij}) \right\rangle, \tag{4}$$

the water-water, water-ion, and ion-ion RDFs were calculated. Integration of $\rho_N g(r)$, where the number density $\rho_N = N/V$, gives the average running total number of neighbors. The number of neighbors in the first solvation shell was computed from the integration of the RDF up to the position of the first local minimum.

To calculate the mass density of the solutions, the total mass of each component calculated from the number of particles (N_x) multiplied by the molar mass (MM_x) was summed and divided by the volume ($\langle V\rangle$) of the simulation cell.

$$\langle \rho \rangle = \frac{N_{\text{water}} MM_{\text{water}} + N_{\text{Na}} MM_{\text{Na}} + N_{\text{Cl}} MM_{\text{Cl}}}{N_A \langle V \rangle}. \tag{5}$$

2. Dynamics

To understand the dynamical nature of the ions in solution, the diffusion coefficient and the residence times of the particles were extracted from the simulation trajectories. Using the mean square displacement, the self-diffusion coefficient D was determined according to the Einstein relation,[97]

$$6Dt = \lim_{t\to\infty} \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle. \tag{6}$$

The mean residence time (MRT) is another dynamical measure we used in determining the validity of the mW-ion model. Typically this has been done using the method developed by Jorgensen *et al.*[98] A recent study has shown, however, that its sensitivity can be questionable, especially regarding the choice of the time during which a pair of particles can still be counted as associated regardless of unsuccessful dissociation events. In this work, we followed the approach of Laage and Hynes[99] to calculate the MRT. Briefly, this was done by recording the distance between the cation and the anion in an ion pair and monitoring the probability of dissociating a CIP as identified by the pair distance reaching a stable state of the SSIP regime ($r_{\text{Na-Cl}} > 4.5$ Å). By constructing a histogram of the times for thousands of dissociation events and fitting this resultant probability to an exponential $\exp(-t/\tau)$, the residence time τ was calculated.

III. RESULTS AND DISCUSSION

A. Solution structure

Capturing and quantifying the tetrahedral character of water around ions was a central objective in the parameter search for developing the mW-ion model. Figure 1 highlights the average tetrahedral character of water in the ionic solutions. The figure presents results from simulations using the atomistic and mW-ion model. Examining the atomistic results, it can be seen that water, when surrounded in the first shell exclusively by water [panel (a)], maintains a high degree of ordering. When one of the four closest neighbors in the first shell is a charged cation [panel (b)], a shift is seen toward a less tetrahedral orientation. Inclusion of a charged anion in the first shell [panel (c)] maintains a strong tetrahedral character. At higher concentrations (data not shown), the solutions generally show less tetrahedral order, though the change is very slight. Comparing the distributions from the TIP4P-Ew and mW-ion simulations, it can be seen that the coarse-grained model qualitatively reproduces the atomistic results with regard to the tetrahedral order.

The RDFs calculated from the atomistic and CG simulations are found in Figs. 2–4 and 7. The data sets labeled as "experimental" in these graphs have two sources: the oxygen-oxygen RDF in Fig. 2 for pure water (topmost solid line) is from Ref. 100, while the other data involving NaCl are from Refs. 68 and 69. In these works, the results of neutron diffraction experiments of water and of NaCl solutions were projected into atomistic Monte Carlo simulations

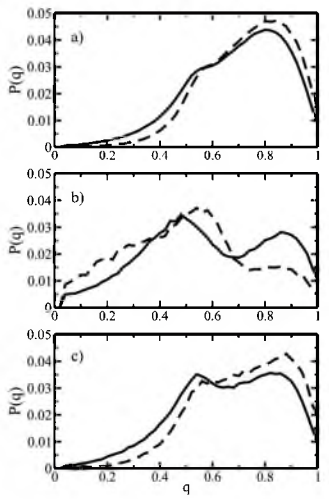


FIG. 1. Probability of observing the tetrahedral order parameter q calculated according to Eq. (1) from configurations of the four closest neighbors around a central water molecule in dilute aqueous NaCl solutions. Results from atomistic and mW-ion simulations correspond to solid and dashed lines, respectively. The panels differentiate between configurations where the four closest neighbors consist of (a) four water molecules, (b) one Na ion and three water molecules, and (c) one Cl ion and three water molecules.

through a technique known as empirical potential structure refinement.[101] Using the densities provided in the references, we integrated the experimental RDFs to give the number of neighbors in the first solvation shell. These results are presented in Table II.

Figure 2 shows the water-water RDF. As stated by Mancinelli *et al.*,[68] the disruption of the hydrogen-bond network of water caused by increasing NaCl concentration affects primarily the second shell of a water molecule. The second peak around 4.5 Å effectively disappears at high concentration, a phenomenon that is reproduced to varying levels of fidelity, by the atomistic and mW-ion models. At this high concentration, the mW-ion RDFs show a peak growing around 3.7 Å that is not witnessed so prevalently for the other two data sets. The position of this peak is due to the ordering of mW water when solvating the ions. Due to the lack of hydrogens, the mW molecule lacks the ability while solvating ions to sample rotational orientations that lead to the more diffuse distribution seen in experiment and atomistic simulations. The water-water first shell solvation decreases steadily across all data sets as the salt concentration is increased, as shown in Table II, due to the integration of the ions into the network of water molecules.

The sodium-water RDF is shown in Fig. 3. There is little change across the different sets of data as the salt concentration increases, both in amplitude and position of the peaks. A slight expansion of the order of 0.3 Å of the second solvation shell of sodium is seen for the mW-ion model. As presented in Table II, the number of waters solvating a sodium ion

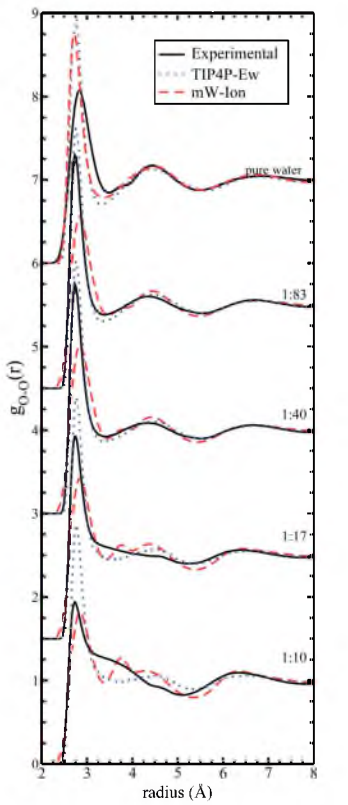


FIG. 2. RDFs for oxygen atoms of water at increasing ion concentration. Experimental values are shown with solid lines: pure water data from Ref. 100; aqueous solutions from Ref 68. Atomistic and mW-ion solutions are represented by dotted and dashed lines, respectively. RDFs at distinct concentrations shifted vertically (+1.5 ordinate units with decreasing ion concentration) for clarity.

decreases through the series of increasing concentration. The drastic decrease in water surrounding the charged sodium ion for the atomistic model partially reveals here some of its weakness; NaCl clustering is prevalent at high concentration as it salts out from the aqueous phase. This observation of unphysical crystallization when modeling 1:1 electrolytes at high concentration has also been reported elsewhere.[91,102]

The water-chloride RDF from the simulations is presented in Fig. 4. The second solvation shells were less defined in the experimental data compared to the atomistic or mW-ion results, though at high concentration the atomistic model shows a similar a "washing-out" effect in the second hydration shell. The coarse-grained results for the chloride-water RDF show a disfavored region between the first and second shell that is overemphasized with regard to the atomistic model and experiment. Nevertheless the positions of

TABLE II. First coordination shell of water molecules around a central water or ion at increasing NaCl concentration.

	Expt.[a]	TIP4P-Ew	mW-ion
	O neighbors of O		
1:83	4.6 ± 0.4[b]	4.4 ± 0.3	4.5 ± 0.4
1:40	4.5 ± 0.3[b]	4.5 ± 0.3	4.4 ± 0.3
1:17	4.4 ± 0.4[b]	4.5 ± 0.3	4.2 ± 0.3
1:10	3.9 ± 0.4[b]	4.2 ± 0.3	3.9 ± 0.3
	O neighbors of Na		
1:83	5.3 ± 0.8[c]	5.5 ± 0.1	5.5 ± 0.1
1:40	5.1 ± 0.9[c]	5.3 ± 0.2	5.3 ± 0.1
1:17	4.6 ± 1.4[c]	4.4 ± 0.2	5.1 ± 0.2
1:10	4.5 ± 1.4[c]	3.6 ± 0.2	4.9 ± 0.1
	O neighbors of Cl		
1:83	6.9 ± 1.0[c]	6.7 ± 0.5	7.1 ± 0.5
1:40	6.8 ± 1.1[c]	6.7 ± 0.7	7.1 ± 0.5
1:17	6.6 ± 1.3[c]	6.3 ± 0.8	7.2 ± 0.8
1:10	6.3 ± 1.3[c]	5.6 ± 0.8	7.2 ± 1.0
	Cl neighbors of Na		
1:83	0.30 ± 0.01[b]	0.45 ± 0.01	0.27 ± 0.01
1:40	0.48 ± 0.02[b]	0.63 ± 0.01	0.42 ± 0.01
1:17	0.68 ± 0.02[b]	1.27 ± 0.02	0.74 ± 0.01
1:10	0.92 ± 0.03[b]	1.90 ± 0.03	0.96 ± 0.02

[a]From Refs. 68 and 69.
[b]Calculated from experimental data of Ref. 68 and 69 provided by the authors.
[c]Reprinted from Ref. 69.

the solvation shell peaks remain in good agreement. The number of water neighbors of chloride (Table II) decreases with increasing concentration for the experimental data and our atomistic simulations, the latter of which shows again evidence of the clustering of ions. For the mW-ion model, however, the number of water neighbors around chloride remains noticeably constant, even slightly increasing as higher concentrations are reached.

After parametrizing the mW-ion model to match closely the RDFs and solvation neighbors, the densities of the resulting solutions were compared to experiment and the atomistic solutions as part of the validation of the model. These densities are plotted in Fig. 5. The results for the mW-ion model show excellent agreement with experimental values, with an average deviation of 0.5%. Using the same potential between CG ions, simulations of NaCl crystal were able to capture the experimental density within 7% despite the fact that the crystal or molten salts were not used in the parametrization of the mW-ion model.

B. Thermodynamics

As with any coarse-grained model, representing all properties of the reference atomistic system using the mW-ion model is impossible.[103] Due to the extremely short-ranged nature of the interactions in our CG system, we expected shortcomings to present themselves with regard to thermodynamic properties. The average total energies calculated from the atomistic and mW-ion simulations are graphed in

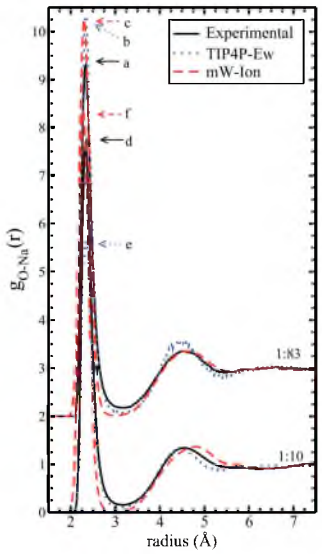


FIG. 3. O–Na RDFs at high and low ion concentrations. Intermediate concentrations vary only slightly and were not graphed for clarity. Experimental values from Ref. 68 are shown with solid lines. Atomistic and mW-ion solutions are represented by dotted and dashed lines, respectively. RDFs at distinct concentrations are shifted vertically for clarity. Also for clarity, the peak apex has been marked for each curve. (a), (b), and (c) correspond to 1:83 experimental, TIP4P-Ew, and mW-ion curves, respectively. (d), (e), and (f) correspond to 1:10 experimental, TIP4P-Ew, and mW-ion curve apices, respectively.

Fig. 6. These correspond to the negative of the energy of vaporization of the solutions per mole of ion pair and show the mW-ion model to follow the same trend as the atomistic one with increasing concentration. The energy of vaporization of pure water is the same in TIP4P-Ew, mW, and experiment.[62,72] On the other hand, the energy of vaporization of the NaCl crystal in mW-ion model is just 35% of the atomistic value. A previous parametrization of the interactions in molten NaCl using potentials that decay within 10 Å accounted for 70% of the atomistic value;[25] aqueous solutions of the salt were not considered in that study. The ratio between the vaporization enthalpies in the mW-ion coarse-grained and atomistic models goes from 90% for the most dilute solution to 70% for the most concentrated solution of this study.

C. Solvent-separated and contact ion pairs in water

The sodium-chloride RDF is presented in Fig. 7. Two types of ion pairs are clearly present from these distributions: contact and solvent-separated ion pairs occurring at around 2.6 and 5 Å, respectively. The formation of CIPs dominates the interaction in experiments, atomistic, and coarse-grained simulations. At low concentration, as seen in Table II, the average numbers of cations around the anions (or vice versa) are 0.30, 0.45 and 0.27, for experiment, atomistic, and mW-

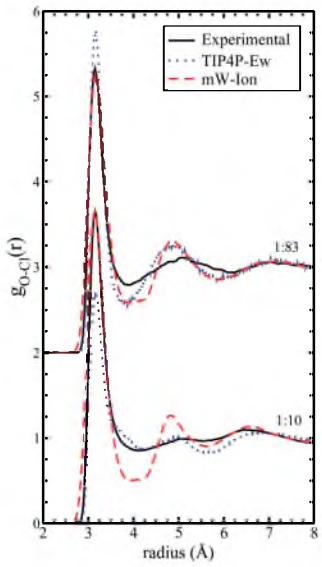


FIG. 4. O–Cl RDFs at high and low NaCl concentrations. Key as in Fig. 2.

ion models, respectively. This coordination increases for both experiment and the mW-ion model (0.92 and 0.96, respectively) at the highest salt concentration to a proportionate degree but shows a drastic increase (1.90) for the atomistic model. The exaggerated CIP peak at higher concentrations for our atomistic simulations is consistent with the overagglomeration of the ions, which will be even further detailed by the residence time data (Fig. 11).

In analyzing the pair formation in the atomistic solutions, it was insightful to monitor the distance between the

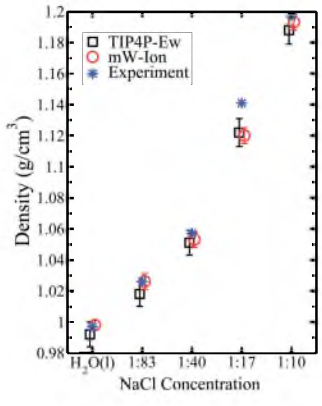


FIG. 5. Densities of NaCl solutions (in g/cm^3) at increasing concentration (ion:water molar ratio) at 298 K. Experimental values from Ref. 106.

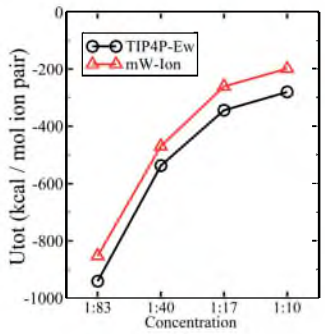


FIG. 6. Average total internal energy (U_{tot}) in kcal/(mol ion pair) at increasing NaCl concentration for atomistic and mW-ion systems that include 1000 water molecule plus ions. Data points are an average of over 50 ns of simulation each.

pairs as a function of the angle formed between them and the solvating water molecule(s) that belong to the first solvation shell of the cation and anion of the pair. Figure 8(b) shows the distribution of these angles and the average distances between the water and ions at each of those angles for the 1:83 NaCl solution. When a water molecule is in the first solvation shell of both a Na and a Cl ion, we calculated the leg distance and angles of the triangle formed [see Fig. 8(a)]. CIPs, seen in the sharp peak of the angle distribution, involve a Na–O–Cl angle around 45°, with fairly constant Na–O and Na–Cl distances. The Cl–O distance, however, varies through the CIP, consistent with a rotation of the bridging hydrogen away from direct contact with the chloride ion. The SSIP peak is more diffuse, giving a distribution of angles centered at around 110°. With constant O–Na and O–Cl distances across the SSIP regime, the optimal values for the other two angles were calculated. These were used in the parametrization of the three-body anion-water-cation interaction to encourage SSIP formation as a balance against the two-body anion-cation interaction that stabilizes the CIP. These parameters are found in Table I for the entries "Cl mW Na" and "Na mW Cl".

To determine if the same mechanism were at play in the mW-ion model, we also monitored the pair distance as a function of the angle in the coarse-grained simulations. A two-dimensional distribution was constructed to show the free energy landscape as a function of the angle and pair distance to further study the differences between the two models. Figure 9 shows the free energy landscape dissociation of a NaCl pair in TIP4P-Ew [panel (a)] and mW [panel (b)] water. The two distributions are strikingly similar: they have a clearly defined minimum corresponding to the CIP configuration and the more diffuse SSIP region separated by a highly disfavored transition state.

The free energy barrier from CIP to the saddle point along the angle-distance coordinate in Fig. 9 is 3.42 ± 0.05 kcal/mol for the atomistic model and

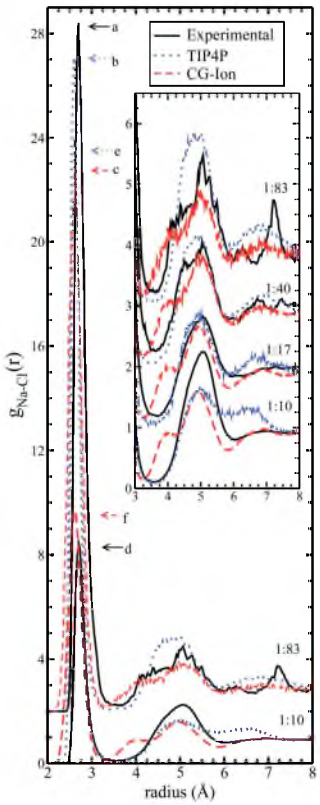


FIG. 7. NaCl RDFs at high and low NaCl concentrations. The inset shows blowup of the same graphs to highlight the second shell (SSIPs). For clarity, the peak apex has been marked for each curve. (a), (b), and (c) correspond to 1:83 experimental, atomistic, and mW-ion curves, respectively. (d), (e), and (f) correspond to 1:10 experimental, atomistic, and mW-ion curve apices, respectively.

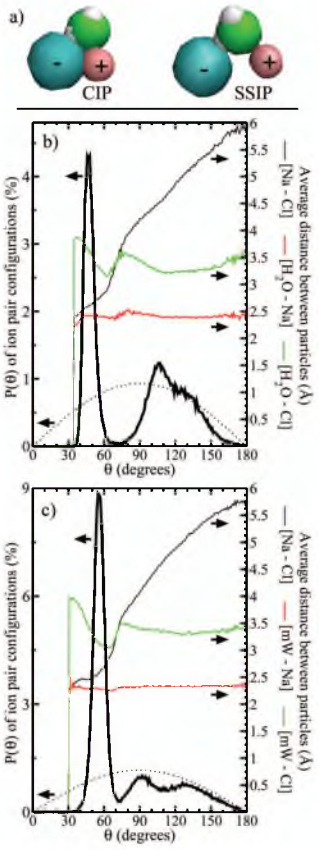


FIG. 8. (a) Snapshots of atomistic simulation configurations of contact and solvent separated ion pairs. (b) Thick solid line gives the probability distribution of angles formed of atomistic (TIP4P-Ew) Cl–H_2O–Na configurations where both Na and Cl reside in the first solvation shell of a water molecule, corresponding to the left axis. The dotted line shows a random distribution for comparison. Also graphed with thin solid lines are the average distances in angstroms between the three combinations of particles for these Cl–H_2O–Na configurations at each angle in the distribution, corresponding to the right abscissa. Of particular note is the differentiation of contact ion and solvent-separated ion pairs at around 45° and 109.5°, respectively. (c) The same as in (b) for the mW-ion model. Note the change in scale for the $P(\theta)$ when comparing (b) and (c). 1:83 NaCl:water concentration was used for both (b) and (c).

2.72 ± 0.03 kcal/mol for mW-ion model. The lower barrier between CIP and SSIP states in the mW-ion model arises from a less defined minimum in the angular distribution (Fig. 8). Without the orientational influence of the protons in the atomistic water model, the CG system is able to sample transition state configurations that are less disfavored. Nevertheless, a strong correlation can be seen between the angular and the pair distance dependence of the two models. Essentially the same mechanism for pair formation and dissociation is observed for the mW-ion and atomistic models.

As an additional measure of the energetic balance between the water-ion interaction and the cation-anion attraction responsible for pair formation, we calculated the association constant of sodium and chloride in the solution. The association constant is defined by

$$K_A = \frac{[\mathrm{NaCl(aq)}]}{[\mathrm{Na^+(aq)}][\mathrm{Cl^-(aq)}]} = \frac{[\mathrm{CIP}]}{([\mathrm{total\ \ salt}] - [\mathrm{CIP}])^2}, \quad (7)$$

where for its calculation in terms of CIP concentration we followed Fennell *et al.*[104] and Chen and Pappu.[102] Although this formulation of the association constant neglects more complicated equilibrium phenomena such as SSIP formation, a comparison can be drawn in this manner with the experimental conductance data.[102] The experimental association

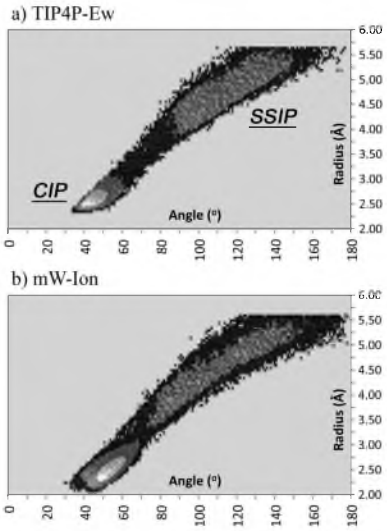


FIG. 9. Potential of mean force for the dissociation of NaCl ion pairs in water as a function of two reaction coordinates: the distance between the ions and the angle formed by the two ions with water molecules that belong to the first coordination shell of the two ions. Panel (a) shows results from atomistic TIP4P-Ew simulations, while panel (b) shows results are from mW-ion simulations. White, gray, and black regions of plot correspond to free energy ranges of 0–1, 1–2.7, and 2.7–4.6 kcal/mol, respectively. The free energies are referred to that of the CIP.

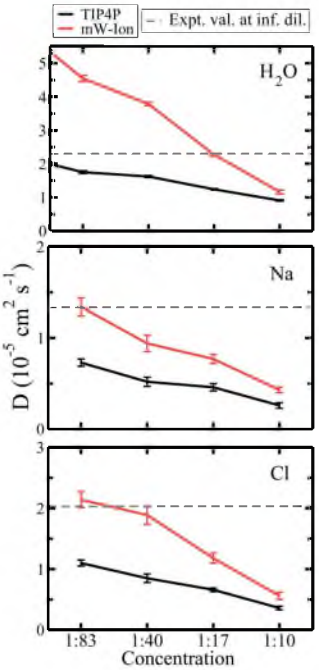


FIG. 10. Self-diffusion coefficients calculated from the mean square displacement of solution particles. Experimental values from Ref. 106.

constant of NaCl has been reported for a $1m$ solution, $K_A = 0.82M^{-1}$.[105] The value of the association constant predicted by atomistic simulations is strongly dependent on the details of the force field. An exhaustive comparison of K_A values computed with three ion force fields and five atomistic water models show K_A for NaCl to range from $0.23M^{-1}$ to $1.2M^{-1}$, with OPLS ions in SPC water giving the closest agreement with experiment, $0.77 \pm 0.06M^{-1}$, among the force fields of that study.[104] We used Eq. (7) to compute K_A from our atomistic and coarse-grained simulations. The value we report was linearly interpolated to $1m$ from the two closest ionic concentrations of our study. We found $K_A = 0.95 \pm 0.07M^{-1}$ for the reference atomistic model and $0.88 \pm 0.07M^{-1}$ for the coarse-grained solution, in very good agreement with the experimental constant $0.82M^{-1}$ extracted from conductimetric studies.[105] This shows the mW-ion model to be capturing well the competing balance between ion solvation and ion-ion attraction.

D. Water/ion dynamics

Using Eq. (6) we calculated the self-diffusion coefficients of ions and water molecules in our simulations. These results are presented in Fig. 10. There is naturally a disparity between those calculated in the atomistic model compared to those in the mW-ion model; coarse-grained particles move typically faster than fully atomistic ones because they evolve in a smoother potential energy surface. For pure water, the diffusion coefficient of the coarse-grained mW liquid is 2.7 times higher than the atomistic. The difference in diffusivity between atomistic and coarse-grained solutions decreases with increasing ionic concentration.

As the ionic concentration is increased, a slowing effect on the diffusion for all species for the atomistic and mW-ion model is seen. Since the coarse-grained system cannot reproduce the exact dynamics of the fully atomistic simulation, we focused on the trends of diffusion due to the increasing concentration to draw the comparison between the two. In Fig. 11 we present the ratios $D_{\text{ion}}/D_{\text{water}}$ for the atomistic and CG simulations along the full concentration range. Three observations emerge: first, the ratio of ion to water diffusion is essentially the same for the atomistic and coarse-grained solutions. Second, the relative diffusion of ions with respect to water does not change significantly with concentration. Third, in the two models the cations are less mobile than the anions, in agreement with experiments.[106]

Since faster diffusion is a characteristic of this coarse-grained system, we expect the MRT of CG particles to be lower than the atomistic. This is confirmed in Fig. 12. As the ion concentration is increased, generally the residence times follow the same trend between the atomistic and the mW-ion model. The exception is the residence of water in the first

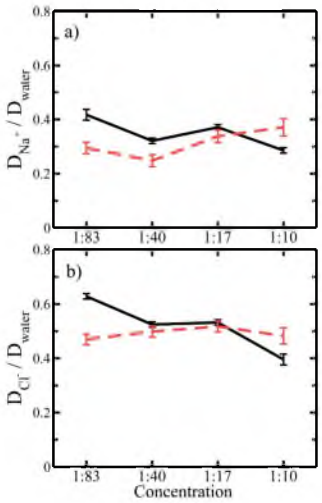


FIG. 11. Ratio of D_{ion}/D_{water} presented in Fig. 10 for the atomistic (black) and mW-ion (dashed red) models along the full concentration range. Panel (a) gives results for the cation, while panel (b) is for the anion.

shell of Na. The mW-ion solution shows the trend we expect as a function of concentration: slower exchange at high concentrations, where the overall diffusion is slower. The atomistic solution shows instead a slight decrease in residence time for water around the ions. This may be related to the clustering of the ions in our atomistic simulations if the residence time at the surface of the aggregate is lower than in solution. Moreover, the aggregation (crystallization) of ions in the atomistic solution is supported by the extremely long Na–Cl residence time at high concentration.

We estimate the dissociation times τ for the NaCl CIP using transition state theory,

$$\tau = \frac{h}{k_B T} e^{\Delta G^{\#}/RT}, \tag{8}$$

where h and k_B are Planck's and Boltzmann's constants and $\Delta G^{\#}$ is the free energy barrier for dissociation computed above. The characteristic times of dissociation predicted using the values of $\Delta G^{\#}$ 3.42 + 0.05 kcal/mol for the atomistic and 2.72 ± 0.03 kcal/mol for mW-ion model, are respectively 52 ± 4 and 16 ± 1 ps for the most dilute solution. These should be compared to the 135 ± 16 and 62 ± 9 ps obtained from the direct analysis of the simulations. The longest time scales observed in the simulations suggest that other coordinates besides the direct cation-anion distance and the minimalistic solvent coordinate, given by the angle between the ions and their common water solvents, play a role in the dynamics of ion pair dissociation. Taking the ratio of dissociation rates for the atomistic and coarse-grained models, as predicted by Eq. (8), shows our simplified reaction coordinate to almost quantitatively capture the ratio garnered from the direct analysis of the trajectories (3.3 + 0.3 versus 2.2 + 0.4, respectively).

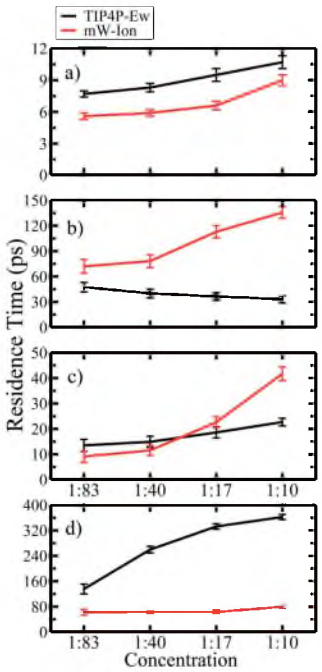


FIG. 12. MRTs of solution species with increasing NaCl concentration; i.e., average duration of first shell solvation (in picoseconds) as calculated in Ref. 99. Panel (a) MRT of water-water, (b) Na-water, (c) Cl-water, and (d) Na–Cl.

E. Benchmarking

By performing simulations using both the mW-ion and the atomistic models, we were able to compare directly the efficiency of the coarse-grained system. Using a system size of 1000 molecules and ions at the lowest NaCl concentration, simulated over 1 ns, we found the mW-ion model with a 5 fs time step to show a 96-fold decrease in the amount of processing time required. To investigate the scaling with the number of particles, we increased the system size by factors of 2, 4, and 16. These experiments showed the mW-ion model to scale slightly less favorably than the atomistic with PPPM; the coarse-grained 16 000-particle system allows a 90-fold increase in computational efficiency. The highest concentration shows a similar scaling, though not as efficient with an average of 65-fold speed-up of coarse-grained versus atomistic model. As a side note, this is a comparison against the PPPM method of computing electrostatic interactions—the Ewald routine, against which comparisons of computational efficiency are normally drawn, is slower than PPPM, with a scaling factor of $N^{3/2}$ versus $N \log N$, respectively.

The mW-ion model scales similar to the number of particles, showing a slight degradation compared to the $N \log N$ scaling of PPPM.

The efficiency increase seen for the mW-ion model is due to three factors: (i) a fivefold increase in the time step compared to the atomistic model, allowed by the softness of the interaction potentials and the lack of fast rotations involving the hydrogen atoms, (ii) a threefold (or less) decrease in the number of particles, and primarily (iii) the use of very short-ranged potentials, with cutoffs at around 5 Å that lead to a significantly reduced neighbor list. The short cutoffs offset the extra cost incurred by the need of computing an extra loop in the calculation of the three-body forces. LAMMPS parallelizes the calculations of the forces through spatial decomposition of the system among the different processors; there is no further optimization of the three-body loops or their parallelization in the code.

IV. CONCLUSIONS

In the previous work, it has been shown that water can be modeled as a single particle without electrostatics or any long-ranged interaction.[62] The monatomic water model, mW, produces "hydrogen-bonded" configurations while using water-water interactions shorter than 4.3 Å. The physical basis of our coarse-graining approach is that the electrostatic attractions and repulsions are shielded in condensed phases of water, leading to effective short-ranged interactions. In spite of the lack of hydrogen atoms and electrostatics, the monatomic water model reproduces the thermodynamic anomalies of water, the structures of the liquid, ice, and low-density amorphous ice, as well as the enthalpy of the phase transitions between them. In this paper we develop an extension of the model for aqueous solutions of ions. The idea behind the mW-ion model is that the water-ion and ion-ion interactions are also highly screened in aqueous solutions and can be represented by very short-ranged potentials.

This may seem a hopeless approach; nevertheless, there have been previous reports in which long-ranged electrostatics were replaced with relatively short-ranged potentials.[26] Izvekov *et al.*[25] modeled molten NaCl in which the ion-ion interactions were represented by a seventh order polynomial decaying at 10–12 Å. The form of the polynomial was similar, but not the same, as the Yukawa potential. The latter models screened Coulomb interactions and has been used to replace electrostatic interactions in the modeling of silica.[20] Electrostatics in silica, as in water, are highly screened and the range of the interactions for them in Refs. 20 and 62 are similar, at around 5 Å. An interesting question is whether the longer ranged ion-ion potentials of Ref. 25 can be combined with the shorter ranged mW potential to produce a model of aqueous solutions. We tried this approach without success; mW water is unable to "shield" the much longer ranged ion-ion interactions resulting in unphysical results. This suggests that the length scale of the ion-ion, ion-water, and water-water interactions have to be comparable to reproduce the solvent effect of water in a coarse-grained model of electrolyte solutions. We found that ion-ion and ion-water interactions that decay in about 5 Å are appropriate to qualitatively reproduce most properties of NaCl solutions discussed in this work.

In our coarse-graining approach, we replace the electrostatic forces by extremely short-ranged pair interactions with three-body orientational effects. Nuances of solvation in an atomistic model, such as the disruption of the tetrahedral ordering of water in the presence of a sodium cation and the more tetrahedral character of water around chloride, are captured by the mW-ion model through the three-body terms of the force field. In moving to a model that uses only short-ranged interactions to effect this structural ordering, our results indicate that the model describes well the physical picture of solvation. The pushing and pulling of all of the other charged species in solution around a central ion effectively cancel out and shield the long-ranged effect of electrostatics. This is exactly the paradigm we wished to represent and exploit by developing the mW-ion model.

The shift to short-ranged interactions raises the question of the model's transferability across a wide range of simulation conditions. We look forward to expanding the model's applicability in describing other ions, as well as evaluating its transferability across a temperature range. With increasing salt concentration, we do not see an evident degradation of the agreement between the experiment or the atomistic models on the one hand, and the mW-ion model on the other. An exception is in the description of the energy of vaporization. This is severely underestimated by the coarse-grained model of pure NaCl even though the model correctly reproduces the experimental vaporization enthalpy of pure water. The agreement of the vaporization enthalpy could be improved within the current paradigm by using the crystal and molten NaCl as part of the force field training set to strike a better balance of attractive and repulsive forces. It is an open question whether this could be achieved without worsening the agreement on the structure.

The parametrization of mW-ion model was primarily based on reproducing the structure of aqueous solutions. The coarse-grained model reproduces quite well the orientational order and radial distribution functions for the different species. Of particular note is the formation of the contact and solvent-separated Na–Cl ion pairs. These presented a challenge for a coarse-grained model without electrostatics, as the SSIP is stabilized by the presence of water molecules between the ions that point the negatively charged oxygen to the cation and positively charged hydrogen to the anion. In the mW-ion model, ions without charges are solvated by water without dipoles, thus the formation of SSIPs has to be built explicitly into the model. To introduce this water-mediated effect in the cation-anion interaction, we use a three-body term in the potential between water, cation, and anion. The very short-ranged nature of this term means that it is only in effect when the water molecule belongs to the first solvation shells of the two ions. This results in a strikingly similar pathway for the dissociation of the atomistic and coarse-grained pairs, illustrated by Figs. 8 and 9. The free energy barriers for CIP dissociation in the coarse-grained and atomistic model are 3.42 ± 0.05 and 2.72 ± 0.03 kcal/mol, respectively. The difference between them is consistent with

the mW-ion CIP having a characteristic time of dissociation that is about two times higher than the atomistic. The dynamics of the mW-ion model is, in general, faster than that of the atomistic model. This is a consequence of the coarse-grained system moving in a smoother potential energy surface. The ratio between the mobilities of water and ions, however, is essentially the same in the atomistic and coarse-grained simulations. This suggests that the short-ranged model accounts well for the relative energies of water-water and water-ion dissociation.

We have developed here a coarse-grained model of aqueous solutions of electrolytes. The mW-ion parameters have been optimized for a particular atomistic reference model of NaCl; nevertheless, we expect that the form of the potential and its length scale, as well as the magnitude of the interaction energies ε will be similar when the model is parametrized for other ions or different atomistic models of the same salt. One distinquishing characteristic of different ions is the number of neighbors in the first shell. This can be tuned with the water-ion terms of the force field. Additionally the association constant, again highly variable between differing pairs of ions and which depends on the balance between direct and water-mediated ion-ion attraction, can be tuned with the strength of the cation-water-anion three-body term. Accurate description of the properties of aqueous solutions of other electrolytes would require a reparametrization of these terms.

Parametrization of the mW-ion model for other electrolytes requires new reference atomistic simulations. Nevertheless, the cost of these reference simulations (only requiring around 1000 molecules) is easily offset by the advantages of using the coarse-grained model in the study of properties that require large simulation cells and long simulation times. One such example of this is in studying supercooled aqueous solutions. A recent study[107] has shown that along with the existence of thermodynamic anomalies, liquid water displays a power-law increase in the correlation length of structure fluctuations in the supercooled regime. To examine these features, a system size of about 250 000 molecules is required, as well as time scales upward of 10 ns per temperature point. To study whether simulations of simple salt solutions can show these anomalies as well, this coarse-grained model can be of benefit since the requirements of system size and simulation times render atomistic simulations ineffectual. In this sense, coarse-grained models, in general, can cover a broad range of needs, one of which is to gain insight into the physics of a system under conditions difficult to access through experiments and/or atomistic simulations (in the previous example because of ice crystallization and computing time, respectively). Also of particular importance is to compare coarse-grained results with those of atomistic models and experiments to gain insight into how much the details matter in determining the properties of electrolyte solutions.

Possible applications of the mW-ion model are in the study of materials and biomolecules. Other studies in our research group have shown the ability to describe well hydrophobic and nonionic hydrophilic interactions with the mW model.[108] Ions, water, and hydrophobic and hydrophilic fragments are the building blocks of biomolecules and synthetic materials. Polyelectrolytes with charged regions and areas of hydrophobicity would benefit from scrutiny with the mW-ion model, we surmise, allowing longer time scales to be accessed while maintaining molecular resolution. Examples of these materials are fuel-cell membranes, proteins, and the nucleic acids DNA and RNA. Caution should be taken, however, in using this short-ranged model in conditions where the anisotropy of the system may lead to charge separation along considerable length scales. One such a case would be the distribution of ions in aqueous solutions at the liquid-vacuum interface. Atomistic simulations suggest that, in some solutions, the cations are more deeply buried in the liquid than the anions.[109] This effect may be difficult to produce with the present parametrization of mW-ion model as there are no restoring (attractive) forces to prevent the separation of the ions after they reached a distance larger than ~5 Å. An ion-density-dependent term could be incorporated to prevent this separation but that would probably negate the advantages of the coarse-grained model. Further studies are needed to validate the use of mW-ion or other short-ranged potentials in the presence of anisotropic forces such as those indicated above.

The success of the mW-ion model in reproducing the solvation structure, orientational and radial distributions, solution densities, formation of CIP and SSIP and their free energy landscape, and the relative mobility of ions to water in the solutions prove that, as concept, a short-ranged force field can be used effectively to model aqueous solutions of salts. Coarse graining of interactions and number of particles in a system always leads to a loss of transferability and representability (the ability to represent different properties with a single model).[103,110] We consider the agreement to be quite good considering the extreme assumptions involved in the development of mW-ion model. For the coarse-grained to be more than an exercise in proving a concept, it has to provide a considerable increase in computational efficiency. Comparisons of simulations with the same number of molecules and spanning the same time using the mW-ion and TIP4P-Ew models show an increase in efficiency by two orders of magnitude for the coarse-grained model. The mW-ion model scales with the number of particles similar to the PPPM method of Ewald summation. These performance gains make mW-ion model a promising coarse-grained tool for exploring longer time scales in complex simulations involving explicit aqueous solutions.

ACKNOWLEDGMENTS

This research was supported by a Collaborative Research in Chemistry grant from the National Science Foundation (Grant No. CHE-0628257). The Center of High Performance Computing at the University of Utah is acknowledged for the allocation of computing time. We thank Thomas E. Cheatham III for his thoughtful review of the manuscript and Maria Antonietta Ricci and Alan Soper for sharing their data on the radial distribution functions of solutions.

[1] B. Honig and A. Nicholls, Science 268, 1144 (1995).
[2] B. Roux, T. Allen, S. Berneche, and W. Im, Q. Rev. Biophys. 37, 15

(2004).
[3] F. Dong, B. Olsen, and N. A. Baker, *In Vitro Techniques*, Biophysical Tools for Biologists Vol. 1 (Elsevier Academic, San Diego, 2008), Vol. 84, p. 843.
[4] E. Zomot, A. Bendahan, M. Quick, Y. F. Zhao, J. A. Javitch, and B. I. Kanner, Nature (London) **449**, 726 (2007).
[5] J. Vieregg, W. Cheng, C. Bustamante, and I. Tinoco, J. Am. Chem. Soc. **129**, 14966 (2007).
[6] F. Sinibaldi, B. D. Howes, G. Smulevich, C. Ciaccio, M. Coletta, and R. Santucci, JBIC, J. Biol. Inorg. Chem. **8**, 663 (2003).
[7] T. E. Cheatham III and M. A. Young, Biopolymers **56**, 232 (2000).
[8] L. Vrbka and P. Jungwirth, J. Phys. Chem. B **110**, 18126 (2006).
[9] A. R. Fersht and V. Daggett, Cell **108**, 573 (2002).
[10] U. R. Pedersen, C. Leidy, P. Westh, and G. H. Peters, Biochim. Biophys. Acta **1758**, 573 (2006).
[11] Y. Duan and P. A. Kollman, Science **282**, 740 (1998).
[12] C. D. Snow, N. Nguyen, V. S. Pande, and M. Gruebele, Nature (London) **420**, 102 (2002).
[13] P. L. Freddolino, F. Liu, M. Gruebele, and K. Schulten, Biophys. J. **94**, L75 (2008).
[14] V. Daggett, Curr. Opin. Struct. Biol. **10**, 160 (2000).
[15] C. M. Dobson, A. Sali, and M. Karplus, Angew. Chem., Int. Ed. **37**, 868 (1998).
[16] P. Ewald, Ann. Phys. **64**, 253 (1921).
[17] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[18] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (IOP, Bristol, 1988).
[19] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).
[20] A. Carre, L. Berthier, J. Horbach, S. Ispas, and W. Kob, J. Chem. Phys. **127**, 114512 (2007).
[21] E. Yakub and C. Ronchi, J. Chem. Phys. **119**, 11556 (2003).
[22] D. Wolf, P. Keblinski, S. R. Phillpot, and J. Eggebrecht, J. Chem. Phys. **110**, 8254 (1999).
[23] P. Demontis, S. Spanu, and G. B. Suffritti, J. Chem. Phys. **114**, 7980 (2001).
[24] D. Zahn, B. Schilling, and S. M. Kast, J. Phys. Chem. B **106**, 10725 (2002).
[25] S. Izvekov, J. M. J. Swanson, and G. A. Voth, J. Phys. Chem. B **112**, 4711 (2008).
[26] Q. Shi, P. Liu, and G. A. Voth, J. Phys. Chem. B **112**, 16230 (2008).
[27] M. Deserno and C. Holm, J. Chem. Phys. **109**, 7678 (1998).
[28] S. J. Plimpton, J. Comput. Phys. **117**, 1 (1995).
[29] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**, 1668 (2005).
[30] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[31] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).
[32] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, J. Phys. Chem. B **105**, 4464 (2001).
[33] R. K. Z. Tan, A. S. Petrov, and S. C. Harvey, J. Chem. Theory Comput. **2**, 529 (2006).
[34] F. Muller-Plathe, ChemPhysChem **3**, 754 (2002).
[35] J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Muller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries, *Advances in Polymer Science: Viscoelasticity, Atomistic Models, Statistical Chemistry* (Springer-Verlag, Heidelberg, 2000).
[36] V. Molinero and W. A. Goddard, J. Phys. Chem. B **108**, 1414 (2004).
[37] W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).
[38] W. G. Noid, P. Liu, Y. Wang, J. W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, J. Chem. Phys. **128**, 244115 (2008).
[39] G. A. Voth, *Coarse-Graining of Condensed Phase and Biomolecular Systems* (CRC, Boca Raton, FL/Taylor & Francis, London, 2009).
[40] Y. Lansac, P. K. Maiti, and M. A. Glaser, Polymer **45**, 3099 (2004).
[41] P. K. Maiti, Y. Lansac, M. A. Glaser, N. A. Clark, and Y. Rouault, Langmuir **18**, 1908 (2002).
[42] S. J. Marrink, A. H. de Vries, and A. E. Mark, J. Phys. Chem. B **108**, 750 (2004).
[43] S. J. Marrink and A. E. Mark, J. Am. Chem. Soc. **125**, 15233 (2003).
[44] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, J. Phys. Chem. B **111**, 7812 (2007).
[45] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. Marrink, J. Chem. Theory Comput. **4**, 819 (2008).
[46] S. O. Nielsen, B. Ensing, V. Ortiz, P. B. Moore, and M. L. Klein, Biophys. J. **88**, 3822 (2005).
[47] S. O. Nielsen, C. F. Lopez, I. Ivanov, P. B. Moore, J. C. Shelley, and M. L. Klein, Biophys. J. **87**, 2107 (2004).
[48] S. O. Nielsen, G. Srinivas, C. F. Lopez, and M. L. Klein, Phys. Rev. Lett. **94**, 228301 (2005).
[49] M. Pickholz, L. Saiz, and M. L. Klein, Biophys. J. **88**, 1524 (2005).
[50] W. Shinoda, R. Devane, and M. L. Klein, Mol. Simul. **33**, 27 (2007).
[51] G. Srinivas, S. O. Nielsen, P. B. Moore, and M. L. Klein, J. Am. Chem. Soc. **128**, 848 (2006).
[52] B. L. Bhargava, R. Devane, M. L. Klein, and S. Balasubramanian, Soft Matter **3**, 1395 (2007).
[53] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein, J. Phys.: Condens. Matter **16**, R481 (2004).
[54] Y. T. Wang, S. Izvekov, T. Y. Yan, and G. A. Voth, J. Phys. Chem. B **110**, 3564 (2006).
[55] A. Y. Shih, A. Arkhipov, P. L. Freddolino, S. G. Sligar, and K. Schulten, J. Phys. Chem. B **111**, 11095 (2007).
[56] P. Poulain, A. Saladin, B. Hartmann, and C. Prevost, J. Comput. Chem. **29**, 2582 (2008).
[57] D. J. Tobias, Curr. Opin. Struct. Biol. **11**, 253 (2001).
[58] B. Hribar, N. T. Southall, V. Vlachy, and K. A. Dill, J. Am. Chem. Soc. **124**, 12302 (2002).
[59] H. L. Tepper and G. A. Voth, J. Chem. Phys. **122**, 124906 (2005).
[60] F. Trovato and V. Tozzini, J. Phys. Chem. B **112**, 13197 (2008).
[61] W. L. Ash, M. R. Zlomislic, E. O. Oloo, and D. P. Tieleman, Biochim. Biophys. Acta **1666**, 158 (2004).
[62] V. Molinero and E. Moore, J. Phys. Chem. B **113**, 4008 (2009).
[63] L. Degreve and F. L. B. da Silva, J. Mol. Liq. **87**, 217 (2000).
[64] L. Degreve, V. M. dePauli, and M. A. Duarte, J. Chem. Phys. **106**, 655 (1997).
[65] D. S. Vieira and L. Degreve, J. Mol. Struct.: THEOCHEM **580**, 127 (2002).
[66] S. Bouazizi, S. Nasr, N. Jaidane, and M. C. Bellissent-Funel, J. Phys. Chem. B **110**, 23515 (2006).
[67] W. M. Cox and J. H. Wolfenden, Proc. R. Soc. London, Ser. A **145**, 475 (1934).
[68] R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper, Phys. Chem. Chem. Phys. **9**, 2959 (2007).
[69] R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper, J. Phys. Chem. B **111**, 13570 (2007).
[70] F. Chen and P. E. Smith, J. Chem. Phys. **126**, 221101 (2007).
[71] A. C. Belch, M. Berkowitz, and J. A. McCammon, J. Am. Chem. Soc. **108**, 1755 (1986).
[72] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, J. Chem. Phys. **120**, 9665 (2004).
[73] J. Aaqvist, J. Phys. Chem. **94**, 8021 (1990).
[74] J. Alejandre and J. P. Hansen, Phys. Rev. E **76**, 061505 (2007).
[75] D. Beglov and B. Roux, J. Chem. Phys. **100**, 9050 (1994).
[76] L. X. Dang, J. Chem. Phys. **96**, 6970 (1992).
[77] L. X. Dang, Chem. Phys. Lett. **227**, 211 (1994).
[78] L. X. Dang, J. Am. Chem. Soc. **117**, 6954 (1995).
[79] L. X. Dang and B. C. Garrett, J. Chem. Phys. **99**, 2972 (1993).
[80] K. P. Jensen and W. L. Jorgensen, J. Chem. Theory Comput. **2**, 1499 (2006).
[81] G. Lamoureux and B. Roux, J. Phys. Chem. B **110**, 3308 (2006).
[82] P. J. Lenart, A. Jusufi, and A. Z. Panagiotopoulos, J. Chem. Phys. **126**, 044509 (2007).
[83] J. N. C. Lopes and A. A. H. Padua, J. Phys. Chem. B **110**, 19586 (2006).
[84] G. H. Peslherbe, B. M. Ladanyi, and J. T. Hynes, Chem. Phys. **258**, 201 (2000).
[85] B. Roux, Biophys. J. **71**, 3177 (1996).
[86] D. E. Smith and L. X. Dang, J. Chem. Phys. **100**, 3757 (1994).
[87] L. S. Sremaniak, L. Perera, and M. L. Berkowitz, Chem. Phys. Lett. **218**, 377 (1994).
[88] T. P. Straatsma and H. J. C. Berendsen, J. Chem. Phys. **89**, 5876 (1988).
[89] O. Teleman and P. Ahlstroem, J. Am. Chem. Soc. **108**, 4333 (1986).
[90] S. Weerasinghe and P. E. Smith, J. Chem. Phys. **119**, 11342 (2003).
[91] I. S. Joung and T. E. Cheatham III, J. Phys. Chem. B **112**, 9020 (2008).
[92] F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).
[93] S. Winstein, E. Clippinger, A. H. Fainberg, and G. C. Robinson, J. Am. Chem. Soc. **76**, 2597 (1954).

[94] H. Sadek and R. M. Fuoss, J. Am. Chem. Soc. **76**, 5897 (1954).
[95] P. L. Chau and A. J. Hardwick, Mol. Phys. **93**, 511 (1998).
[96] J. R. Errington and P. G. Debenedetti, Nature (London) **409**, 318 (2001).
[97] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 1st ed. (Clarendon, Oxford, 1987).
[98] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[99] D. Laage and J. T. Hynes, J. Phys. Chem. B **112**, 7697 (2008).
[100] A. K. Soper, J. Phys.: Condens. Matter **19**, 335206 (2007).
[101] A. K. Soper, Chem. Phys. **202**, 295 (1996).
[102] A. A. Chen and R. V. Pappu, J. Phys. Chem. B **111**, 6469 (2007).
[103] M. E. Johnson, T. Head-Gordon, and A. A. Louis, J. Chem. Phys. **126**, 144509 (2007).
[104] C. J. Fennell, A. Bizjak, V. Vlachy, and K. A. Dill, J. Phys. Chem. B **113**, 6782 (2009).
[105] R. M. Fuoss, Proc. Natl. Acad. Sci. U.S.A. **77**, 34 (1980).
[106] *CRC Handbook of Chemistry and Physics*, edited by D. R. Lide (CRC, Boca Raton, FL/Taylor & Francis, Boca Raton, FL, 2009).
[107] E. B. Moore and V. Molinero, J. Chem. Phys. **130**, 244505 (2009).
[108] L. Le and V. Molinero (unpublished); N. Kastelowitz and V. Molinero (unpublished); L. Xu and V. Molinero (unpublished).
[109] P. Jungwirth and D. J. Tobias, J. Phys. Chem. B **106**, 6361 (2002).
[110] A. A. Louis, J. Phys.: Condens. Matter **14**, 9187 (2002).

CHAPTER 3

A COARSE-GRAINED MODEL OF DNA WITH EXPLICIT SOLVATION BY WATER AND IONS

This chapter was reproduced from the published paper with permission from DeMille, R. C., Cheatham, T. E. III, & Molinero, V. (2011) *The journal of physical chemistry B* **115**, 132-142.

A Coarse-Grained Model of DNA with Explicit Solvation by Water and Ions

Robert C. DeMille,[†] Thomas E. Cheatham III,[‡] and Valeria Molinero*,[†]

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112-0850, United States, and Department of Medicinal Chemistry and Department of Pharmaceutics and Pharmaceutical Chemistry, College of Pharmacy, University of Utah, 2000 South 30 East, Skaggs Hall 201, Salt Lake City, Utah 84112, United States

Received: July 27, 2010; Revised Manuscript Received: November 22, 2010

Solvation by water and ions has been shown to be vitally important for biological molecules, yet fully atomistic simulations of large biomolecules remain a challenge due to their high computational cost. The effect of solvation is the most pronounced in polyelectrolytes, of which DNA is a paradigmatic example. Coarse-grained (CG) representations have been developed to model the essential physics of the DNA molecule, yet almost without exception, these models replace the water and ions by implicit solvation in order to significantly reduce the computational expense. This work introduces the first coarse-grained model of DNA solvated explicitly with water and ions. To this end, we combined two established CG models; the recently developed mW-ion model [DeMille, R. C.; Molinero, V. *J. Chem. Phys.* **2009**, *131*, 034107], which reproduces the structure of aqueous ionic solutions without electrostatic interactions, was coupled to the three-sites-per-nucleotide (3SPN) CG model of DNA [Knotts, T. A., IV; et al. *J. Chem. Phys.* **2007**, *126*, 084901]. Using atomistic simulations of d(CGCGAATTCGCG)$_2$ as a reference, we optimized the coarse-grained interactions between DNA and solvent to reproduce the solvation structure of water and ions around CG DNA. The resulting coarse-grained model of DNA explicitly solvated by ions and water (mW/3SPN-DNA) exhibits base-pair specificity and ion-condensation effects and it is 2 orders of magnitude computationally more efficient than atomistic models. We describe the parametrization strategy and offer insight into how other CG models may be combined with a coarse-grained solvent model such as mW-ion.

1. Introduction

The DNA molecule has been the subject of many molecular simulation studies in an attempt to understand its varied physical phenomena at atomic resolution.[1−11] This includes the development of coarse-grained (CG) models that improve the efficiency of DNA modeling by using different levels of coarsening to focus on different aspects of the physics of the DNA molecule, such as plasmid supercoiling,[12] protein docking,[13] persistence length,[14,15] and denaturation.[16−18] These CG models neglect explicit solvation for the sake of computational efficiency, choosing instead to implicitly define the interaction of DNA with water and ions. The structure and dynamics of DNA, however, are strongly influenced by solvation and ions, and the exclusion of specific hydration and ionic interactions has been shown to have deleterious effects on the structure of DNA.[19−21] Significant advances have been made in coupling a particle-based implicit solvent with atomistic DNA, resulting in about 5 times increase in efficiency with respect to atomistic simulations.[22] Another CG DNA model focusing on helix formation includes a similar explicit solvent but neglects features such as base-specificity and the counterions.[23] Due to their exclusion of explicit water and/or ions, these various models suffer from limited applicability for studying the structure of DNA in solution and processes that involve solvation/desolvation of the DNA.

Being a polyelectrolyte, the polyanionic nature of DNA and the closeness between the phosphate groups along the strand lead to counterion association. Described theoretically by Manning,[24,25] counterion condensation has been subject to many experimental and simulation studies. For a thorough review, we refer the reader to refs 19 and 26−30 and references therein. The attraction of the negatively charged phosphate moieties of DNA to the cationic counterions in solution involves strong electrostatic interactions. When solvated by water and cations, the long-range influence of the polyanion is effectively diminished due to the shielding of the charges. We hypothesize that the physical picture of the solvated, counterion-condensed DNA can be captured with a coarse-grained model with short-ranged interactions. Along these lines, we have recently developed a coarse-grained model of water and ions, the mW-ion model,[31] where each water molecule and ion (Na and Cl in ref 31) is represented by a single particle interacting through very short-ranged potentials. The model is based on the monatomic water model mW,[32] which reproduces the structure and anomalies of liquid water, ice, clathrates, and low-density amorphous ice and the phase transformations between them.[32−37] The key for the success of the mW model is its description of "hydrogen-bonded" structures through the use of a combination of two- and three-body potentials that encourage tetrahedral configurations of the water molecules.[32] The mW-ion model uses a similar strategy, plus the introduction of Yukawa potentials to modulate the repulsion between same-charge ions, to reproduce the density, radial and angular distribution functions, equilibrium between contact and solvent-separated ions pairs, and relative mobility of ions with respect to water of NaCl aqueous solutions up to a concentration of at least 5 M.[31] There are no electrostatics

* Corresponding author. E-mail: Valeria.Molinero@utah.edu.
[†] Department of Chemistry.
[‡] Department of Medicinal Chemistry and Department of Pharmaceutics and Pharmaceutical Chemistry.

10.1021/jp107028n
Published on Web 12/14/2010

or other long-ranged interactions in the mW and mW-ion models; chargeless ions interact with mW water and other ions via short-ranged potentials. We refer the reader to ref 31 for the full details of mW-ion, its parametrization, and its unique ability among coarse-grained models to replicate solution structure without hydrogen atoms and electrostatic interactions. Of particular note is the computational efficiency of the mW-ion model; molecular dynamics simulations with the mW-ion model are 2 orders of magnitude faster than with rigid atomistic water and ion models with Ewald sums. The increased efficiency is due to the lack of hydrogen atoms, increased time step, and the short range (between 4.3 and 7 Å) of the interactions.[31,32]

A natural progression and challenge is to interface the mW-ion model to coarse-grained models of biological molecules. Atomistic simulations, although becoming more negotiable with more efficient hardware and software implementations, are still intractable for long simulations (μs−ms) involving explicit solvation. On the other hand, typical CG implementations involving implicit solvation of DNA are too simplified to capture important interactions mediated by specific ion and water interactions, such as protein recognition and drug interactions. In the present work, we combine the mW-ion model with the well-performing and efficient 3SPN.0 coarse-grained DNA model and force field of Knotts and co-workers[17] to produce a coarse-grained model of nucleic acids with an explicit representation of water and ions that we name mW/3SPN-DNA. The 3SPN model reproduces several key features characteristic of DNA, including (i) sequence specificity by the incorporation of four distinct nucleotides; (ii) B-form helicity, structure, and stability through bonded and nonbonded interactions among phosphate, sugar, and base moieties; and (iii) semiquantitative reproduction of melting behavior. An improved version of this model, the 3SPN.1 DNA model, added an attraction between sugar moieties in opposing strands, achieving quantitative melting and rehybridization behavior when compared to the experiments.[38−40] The attraction between sugars is intended to emulate the effect of the solvent; thus, we base our parametrization on the 3SPN.0 model to avoid double counting of the effect of the solvent by its introduction in implicit and explicit terms of the potential. We note, however, that the methodology we present is general and could be extended to more recent parametrizations of this or other DNA models with comparable degrees of coarse-graining.

The central aim of this paper is to develop and validate the mW/3SPN-DNA coarse-grained model of DNA with explicit solvation by coarse-grained water and ions. A main question is whether the solvation structure of DNA can be accurately reproduced with a coarse-grained model without electrostatics. Here we show that, despite the replacement of long-ranged with short-ranged interactions and the coarse representation of water and DNA, the integrated mW/3SPN-DNA model is able to reproduce the solvation structure seen in atomistic simulations and the relative residence times of water and ions on the DNA moieties. The results suggest that the approach of this study could be extended to develop accurate solvated coarse-grained models of other complex biological systems, such as membranes, proteins, and RNA.

The paper is organized as follows: section 2 presents the models, the parametrization strategy, and the simulation methods; section 3 presents the results, with particular emphasis on DNA structure and stability, solvation structure, and solvation dynamics. After a brief discussion of the computational efficiency of the mW/3SPN-DNA model, the conclusions are presented in section 4.

2. Models, Parmeterization, and Simulation Methods

A. Models. ***Coarse-Grained DNA.*** The 3SPN force field involves bonded and nonbonded terms, corresponding to eqs 2a−2g of ref 17, including an implicit representation of solvent through Debye−Hückel Coulombic screening of the interactions between phosphates. For further insight into the nuances of the force field, including the Gō-type construction of the nonbonded interactions, we refer the reader to the original publication. In our implementation of the 3SPN model, we omit the shielded Coulomb interaction between phosphate sites (eq 2g of ref 17) as we explicitly include the ions and water molecules. Our initial tests with the original 3SPN model involving implicit solvation showed the Coulombic interaction to contribute only approximately 1% to the total energy of the system, and exclusion of the Coulombic interaction appears to have minimal effect on the structure of the DNA duplex, especially at high NaCl concentrations (∼1 M or greater).

We have used the classic Dickerson DNA dodecamer duplex d(CGCGAATTCGCG)$_2$ for the atomistic and coarse-grained simulations of this work. The middle panel of Figure 1 shows the CG sites of the 3SPN model superimposed on an all-atomistic representation of the DNA dodecamer. We followed the prescription for building molecules with standard coordinates, gleaned from the standard form the B-form helix of DNA (PDB ID: 1BNA),[41] previously described in ref 17. We note that this standard B-form structure used as reference for the parametrization of the 3SPN.0 model differs from the actual crystallographic structure,[42] which presents a modulation of the distances between the two strands along the minor groove, while this distance is constant in the standard model. The implications of this difference are discussed in section 3.A below.

Coarse-Grained Water and Ions. The monatomic water mW model[32] and its extension to ionic solutions, the mW-ion model,[31] were recently developed to study the behavior of aqueous ionic solutions at length and time scales that are not easily accessed by atomistic simulations. The mW and mW-ion use the short-ranged interactions of the Stillinger−Weber (SW) potential[43] and rely on the interplay between two-body attraction terms (ϕ_2), which favor high coordination, and three-body repulsion terms (ϕ_3), which encourage tetrahedral configurations, to reproduce the structure of liquid water and the ion solvation structure found in atomistic simulations and experiments. The functional form of the SW potential is shown in eq 1,

$$
\begin{aligned}
E &= \sum_i \sum_{j>i} \phi_2(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{k>j} \phi_3(r_{ij}, r_{ik}, \theta_{ijk}) \\
\phi_2(r_{ij}) &= A\varepsilon_{ij}\left[B\left(\frac{\sigma_{ij}}{r_{ij}}\right)^p - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^q\right] \exp\left(\frac{\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right) \\
\phi_3(r_{ij}, r_{ik}, \theta_{ijk}) &= \lambda_{ijk}\varepsilon_{ijk}[\cos\theta_{ijk} - \cos\theta_{0ijk}]^2 \times \\
&\quad \exp\left(\frac{\gamma\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}}\right) \exp\left(\frac{\gamma\sigma_{ik}}{r_{ik} - a_{ik}\sigma_{ik}}\right)
\end{aligned}
\tag{1}
$$

where r_{ij} is the distance between particles i and j and θ_{ijk} is the angle subtended by the vectors between the positions of the $i-j$ and $i-k$ pairs of particles. The constants that define the potential are $A = 7.049\,556\,277$, $B = 0.602\,224\,558\,4$, $p = 4$, $q = 0$, and $\gamma = 1.2$. The adjustable parameters are the characteristic size, σ; the interaction strength, ε; the strength of the tetrahedral interactions, λ; the cutoff parameter, a; and the preferred angle, θ_o. The water−water parameters were adjusted in ref 32 to reproduce the experimental temperature of melting,

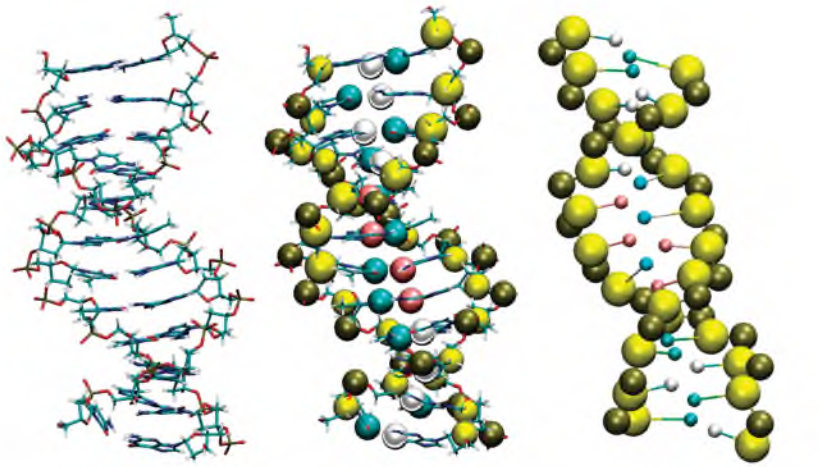


Figure 1. Representation of the d(CGCGAATTCGCG)$_2$ duplex for the atomistic model (left) and superimposition of 3SPN CG sites on top of the atomistic model (middle). (Right) A snapshot from a simulation of mW/3SPN-DNA used in this study, where the sizes of the bases have been reduced for clarity. Particle colors correspond to the following CG sites: G, white; C, dark blue; A, pink; T, light blue; P, olive green; and S, yellow.

density, and enthalpy of vaporization of liquid water. The water−ions and ion−ion parameters were tuned in ref 31 to reproduce the microscopic structure of NaCl−water solutions.

An additional shielded Coulomb Yukawa potential[44] is added to the interactions between cation pairs, as well as anion pairs,

$$E = C \exp(-\kappa r_{ij}) r_{ij}^{-1} \quad r_{ij} < r_{\text{cutoff}} \tag{2}$$

where the magnitude of C correlates with the strength of the shielded-Coulomb interaction and its sign indicates whether it is attractive or repulsive, κ is the screening length, and r_{cutoff} is a distance beyond which the contribution of this term to the energy is negligible. The values of r_{cutoff} in the mW-ion model do not exceed 7 Å, while the interactions described by eq 1 are even shorter ranged. The full details of the force field and the parameters used in the mW-ion model can be found in ref 31; these were used for water and NaCl in the present study.

Explicit Solvation of CG DNA. A combination of the two CG force fields requires careful attention to the cross interactions to provide the necessary balance between the stability of the DNA molecule and its solvation by water and ions. The model system used for parametrization of the solvation forces was a B-DNA model of the d(CGCGAATTCGCG)$_2$ duplex including 22 net-neutralizing Na "counterions" and an aqueous solution of 278 NaCl pairs in 15 000 mW molecules (representing ~1 M NaCl). Simulations at a lower concentration of 200 mM NaCl were also performed, consisting of the DNA duplex solvated by 57 NaCl pairs in ~15 400 waters. The interaction of water and ions with phosphate is described by eqs 1 and 2. The interaction of water and ions with the sugar and base sites involves a simplified form for the SW potential (eq 1) having only two-body interactions (i.e., $\lambda = 0$). The repulsive force between phosphate and chloride was additionally handled with a Yukawa potential (eq 2). A benefit of the CG model is that longer molecular dynamics integration time steps relative to all-atom simulations are possible while still conserving energy. The short-ranged nature of all interactions in the model, the decreased number of particles, and the longer time steps (5 fs) are responsible for its ~90 times increased efficiency with respect to the atomistic model (see section 3.D). All interactions go effectively to zero between 4 and 9 Å. The full set of parameters used in this study for the mW/3SPN-DNA model is provided in Tables S1 and S2 of the Supporting Information.

Atomistic Models. All-atom (AA) simulations of explicitly solvated DNA were used as a reference for the parametrization of the mW/3SPN-DNA model. The initial AA model of d(CGCGAATTCGCG)$_2$ was built from the first model structure from the high-resolution NMR structures of Tjandra and co-workers (PDB ID: 1DUF),[45] and was solvated with TIP4P-EW water molecules[46] surrounding the DNA 15 Å in each direction into a truncated octahedral periodic box. Charge-neutralizing Na^+ ions were added, followed by the addition of 168 Na^+ and Cl^- ions for an approximate added salt concentration of ~1 M[47] (33 NaCl pairs were used for a 200 mM salt solution). The position of each of the ions was randomized by swapping positions with random water molecules such that no ion was within 6 Å of another ion nor within 4 Å of the DNA (neglecting distance checks for ion−ion interaction across the periodic boundaries). When analyzing the AA trajectories, each was first mapped frame by frame into a CG trajectory using the same rules for assigning the CG sites prescribed by the 3SPN model: P at center of mass of the PO_4 group, sugar at center of mass of the deoxyribose unit, pyrimidine base sites C or T at the N3 position, and G and A at the N1 position for the purine bases. The AA water molecules were coarsened similarly by removing the hydrogen atoms.

B. Simulation Methods. The simulations of the coarse-grained mW/3SPN-DNA model were performed using the LAMMPS[48] molecular dynamics simulation package under isobaric−isothermal conditions (NpT) at $p = 1$ atm and at $T = 25$ °C, except when otherwise indicated. The Nose−Hoover thermostat and barostat were used to control the temperature and pressure of the system, with time constants of 0.5 and 5.0

ps, respectively. We extended LAMMPS in order to incorporate the force field of the 3SPN.0 model. NaCl solutions of water, 0.2 or 1.0 M, were first equilibrated for 10 ns. Removing all water and ions in a cylinder centered in the simulation cell with rough dimensions 1.8 nm diameter by 4 nm height allowed the insertion of the dodecamer DNA strand. Equilibration of the system was then performed over 10 ns to allow the ions and water to completely solvate the mW/3SPN-DNA. The equations of motion were integrated with the velocity Verlet algorithm with a time step of 5 fs. Typical production runs were on the order of 100 ns.

The atomistic models of solvated DNA was simulated using a time step of 2 fs at a temperature of 300 K and 1 atm pressure using the Amber 10 package of molecular simulation programs.[49] The temperature was regulated with a Berendsen thermostat with a 5 ps decay constant.[50] Particle mesh Ewald[51] was used for electrostatic interactions with a cutoff of 9 Å and automated builds in a 1 Å buffered pairlist, cubic spline interpolation and an ~1 Å FFT grid. Trajectories comprising approximately 150 ns of simulation time for each ionic concentration were used for the data analysis presented here.

C. Parameterization. Three properties calculated from the AA trajectories with 1 M NaCl were used as targets for our parametrization efforts. We calculated radial distributions (rdf) of water and ions around the DNA CG sites, as well as the number of neighbors in the first shell in order to quantify the solvation structure around the DNA molecule. Additionally, as a measure of the strength of interactions between the CG sites with the mW-ion particles, residence times were calculated to determine the length of time that particles were associated with one another in their first solvation shells. These properties were calculated as we previously reported in the development and validation of the mW-ion force field.[31] The rdf, number of neighbors, and residence times were used as parametrization targets referenced against atomistic simulations with 1 M NaCl. The 1 M solution was chosen because it has a relatively high concentration of electrolyte, the regime for which the effective interactions between ions are expected to be shorter ranged and thus more amenable for its modeling through short-range potentials, while it is not high enough to shift the equilibrium between A- and B-forms of DNA. As molecules in CG models diffuse faster than in atomistic simulations, we did not attempt to reproduce the actual residence times for the ions and water in the first solvation shell of DNA but the *relative* times with respect to the mobility of water model in the system. In ref 31 we have shown that the ratio of diffusivity of ions with respect to water is well-reproduced by the mW-ion model, although the diffusion coefficients themselves are almost 3 times larger than in the atomistic simulations.

The primary method by which the force field was tuned to produce a better fit to AA data was in the adjustment of the parameters σ, ε, λ, and a of the solvent−DNA interaction potential. Upon examination of the atomistic data, we hypothesized the interaction of the water and ions with phosphate to play the most important role in the solvation forces. We initially focused on matching the solvation structure around the phosphate groups, leaving the interaction with the sugar and base groups for fine-tuning. Adjusting σ_{mW-P} and σ_{Na-P} in particular allowed the solvation shells of water and ions around mW/3SPN-DNA to closely align with the target data. Additional narrowing of the solvation shells, found to be necessary, for example, in the case of the interaction of mW with the CG phosphate moiety in order to match the atomistic data, was accomplished by lowering the value of the a parameter, directly affecting the length scale of the potential.

Adjustment of ε helped tune the number of neighbors in the solvation shells and allowed the residence time to more closely match the target. Again, we focused on the interaction between the solvating mW and ions with phosphate in our initial parametrization. Using sequential modifications to the value of ε_{mW-P} allowed the number of water neighbors surrounding mW/3SPN-DNA to be matched to the atomistic data. Generally, an increase in ε leads to an increase in the number of neighbors, as well as an increase in the residence time. This is, however, by no means a linear process; changing the parameters of one particular interaction often also influences others. The parameters of Tables S1 and S2 of the Supporting Information produce our best fit to the target data, though do not exhaustively explore the approximately 1000 tunable parameters available in the cross interactions of the water and ions with the DNA CG moieties.

D. Validation Measures. To assess the validity of the mW/3SPN-DNA model compared to atomistic data and to the implicitly solvated 3SPN.0 DNA model, we analyzed the structural and thermal stability of the DNA molecule, as well as the profile of water and ions surrounding its central axis for DNA in aqueous solutions of two ionic contents, 0.2 and 1 M NaCl. The structural stability of the DNA molecule is addressed in further detail below in section 3.A. The cylindrical distribution of water and ions around the curved central axis of DNA was calculated in the following fashion in order to show the short- and long-ranged structure of solvation of DNA. Disregarding the two terminal base pairs on either side of the DNA duplex of our study, we connected lines between consecutive base pairs. Next, the center of this line was translated to the origin, and the frame of reference rotated to set the line's terminal points at $\pm z_t$ (=1/2 the line distance), propagating the translation and rotation for the entire system of particles while maintaining the requirements of the periodic boundary conditions. For this arrangement, performed with each subsequent base pair for the entire trajectory, the radial distance $r_i = (x_i^2 + y_i^2)^{1/2}$ for all particles i with coordinates $|z| < z_t$ was binned every 0.1 Å to give a distribution of DNA particles, water, Na, and Cl around the actual central axis along the DNA oligomer. The data presented in section 3.B corresponds to averages over 100 ns simulation trajectories.

3. Results and Discussion

In developing an explicitly solvated model for DNA with CG water and ions, a critical aim was to retain the ability to accurately model the structure and thermal stability of the DNA molecule. Structure and stability results are compared to the explicitly solvated atomistic simulations and the implicit-solvation 3SPN.0 coarse-grained model. Additionally, the solvation structure of water and ions around the DNA molecule was replicated with the mW/3SPN-DNA model. The distribution of water and ions surrounding DNA, as well as a measure of the solvation dynamics—the residence times—further quantify the agreement between mW/3SPN-DNA and reference atomistic simulations. In what follows, we present the development and validation of the explicitly solvated coarse-grained DNA model in three subsections concerning (A) DNA structure and stability, (B) the solvation structure around DNA by water and ions, and (C) the residence times of water and ions on the DNA moieties.

A. DNA Structure and Stability. To measure the ability of mW/3SPN-DNA to maintain the helical structure of DNA, we compared the end-to-end distance of the d(CGCGAATTCGCG)$_2$ duplex with the atomistic simulations as well as the

TABLE 1: DNA Structural Stability of the Three Models Used in This Study for 1 M NaCl (data averaged 100 ns trajectories)

atomistic	3SPN.0	mW/3SPN-DNA
End-to-End Distance (Å)		
29.8 ± 0.1	33.4 ± 0.3	33.7 ± 0.2
Rmsd (Å) from Average Structure		
1.3 ± 0.3	2.8 ± 0.8	2.2 ± 0.4
Rmsd (Å) from Standard B-Form		
2.1 ± 0.4	4.6 ± 1.1	4.0 ± 0.5

3SPN.0 model. For this measurement, as well as any subsequent analysis of the DNA solvation structure, the first base pair on either end of the duplex was disregarded to preclude potential artifacts arising from the different solvation environment experienced at the termini (i.e., fraying), in accordance with previous studies.[19] Monitoring this end-to-end distance between the first guanine bases at the 5′ ends of the complementary strands gives a good comparison across the different models regarding the helical structure and compactness of the DNA duplex. This data is presented in Table 1 for the DNA in 1 M NaCl and in Table S3 of the Supporting Information for the 0.2 M NaCl solution. The standard structure the B-form duplex of DNA gives reference coordinates for the B-form Dickerson dodecamer[41] (which both the mW/3SPN-DNA and 3SPN.0 model use to build starting DNA structures) that yields a reference separation of ~30.5 Å for these two G bases. The atomistic model maintains a slightly more compact structure with an end-to-end distance of 29.8 ± 0.1 Å in the 1 M and 29.7 ± 0.1 Å in the 0.2 M solution. The 3SPN.0 model shows a more elongated structure, averaging 33.4 ± 0.3 and 33.8 ± 0.2 Å, respectively. The slightly more elongated structure may be related to the small shifts in the backbone angles and dihedrals with respect to the reference standard B-form structure (Figures S1 and S2 of the Supporting Information). The mW/3SPN-DNA model replicates 3SPN.0 very closely (33.7 ± 0.2 Å in 1 M and 34.2 ± 0.2 Å in 0.2 M), showing that explicit solvation has no deleterious effect on the structure of the 3SPN DNA duplex. This suggests that the mW/3SPN-DNA model captures the subtle balance between the solvation forces of the duplex and the strength of DNA's stiffness inherent to the 3SPN.0 model.

To further compare the structure and stability of the DNA helix between the atomistic and coarse-grained models, we calculated the root-mean-squared deviations (rmsd) of representative 100 ns trajectories from both the trajectory-averaged structure and the standard B-form structure for each of the three models. The analysis was done in VMD 1.8.6,[52] and the rmsd data of the DNA helix can be found in Tables 1 and S3. The atomistic model shows the tighter fluctuations around its average structure, with an rmsd of about 1 Å. The coarse-grained models are softer and allow larger fluctuations, especially the 3SPN.0 model (~3 Å rmsd). Addition of the explicit solvent with the mW/3SPN-DNA model damps these structural fluctuations, leading to an rmsd intermediate between the one of the fully atomistic model and that of the coarse-grained model with implicit solvation.

B-DNA presents major and minor grooves along the helix, where the backbones of the two strands are farther and closer apart, respectively. The upper panel of Figure 2 presents the minor groove average distance between the two strands along the dodecamer for the atomistic, mW/3SPN-DNA, and 3SPN models in 1 M NaCl, as well as the standard reference structure used to build and parametrize the coarse-grained models. We

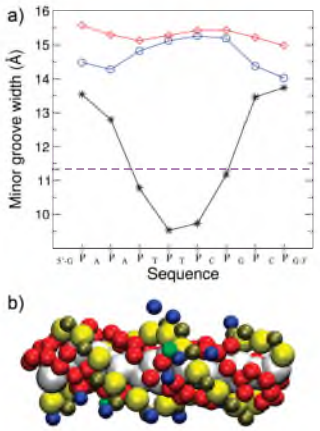


Figure 2. (a) Minor groove distances along the DNA duplex. Atomistic, 3SPN.0, and mW/3SPN-DNA data are symbolized by stars, diamonds, and circles, respectively. The constant minor groove distances of the reference model are shown by a dashed line. The coarse-grained models do not reproduce the thinning of the minor groove observed in the atomistic model and the experiment, presenting an almost constant distance similar to the reference B-DNA form structure used in the 3SPN.0 parametrization. (b) Snapshot of the solvated DNA dodecamer showing the water molecules (red) within 6 Å of both strands and Na (blue) and Cl (green) ions within 6 Å of either strand. DNA moieties are colored as in Figure 1, except for all bases are shown here in white for clarity.

evaluated the groove widths along the sequence by calculating the distance between a phosphate in the tagged strand and the closest phosphate in the other strand across the groove, averaging these quantities over the simulation trajectories at 25 °C (except for the reference B-form structure, which is a single configuration). Experimental results[42,53] and the atomistic simulations show a narrowing of the minor groove in the AATT regime of the dodecamer. This is not reproduced by the reference structure, which presents a constant minor groove width of 11.3 Å along the sequence of the duplex. Not surprisingly, a similarly constant trend is observed for the 3SPN.0 or mW/3SPN-DNA models constructed from the reference B-form structure. The coarse-grained model with implicit solvation shows a fairly constant distance along the sequence, with a value larger than the AA simulations and reference model. The larger distance may be due, in part, to the extension of the end-to-end distances in the coarse-grained DNA. The introduction of explicit water and ions into mW/3SPN-DNA decreases the phosphate−phosphate separation but also does not reproduce the sequence-specific narrowing of the AATT region. Different from the experiment and atomistic models, the minor and major grooves have almost the same width in the coarse-grained models. The resolution of the CG models is sufficient to differentiate the major and minor grooves; thus, the presence of grooves with similar widths is likely due to the form of the force field. Figure S2 of the Supporting Information presents distributions of the backbone dihedral angles for the atomistic and coarse-grained models that show that the CG DNA has unimodal distribution of dihedral angles (as commanded by the dihedral term in the 3SPN.0 potential), while the AA distribution of some dihedral angles is bimodal. The modulation of the width of the minor groove in the atomistic model is correlated to the number of water molecules (one and two) between the two strands along

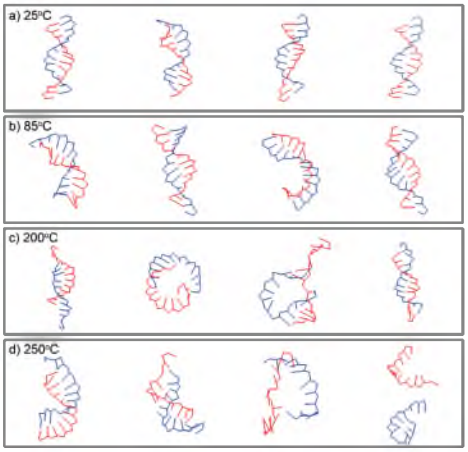


Figure 3. Representative snapshots from 100 ns mW/3SPN-DNA simulations of the Dickerson DNA dodecamer in 1 M NaCl at increasing temperatures. All simulations start from the equilibrated structure shown in the first snapshot of panel a. Note the presence of distorted structures at 85 °C (panel b) and the unzipping of the duplex at 200 °C (panel c), which, however, hybridizes back to a structure similar to its native form. At 250 °C (panel d) irreversible melting is witnessed after 10 ns. Complementary strands are colored in blue and red for clarity.

the sequence. The constant width of the minor groove in the mW/3SPN-DNA model results in about two water molecules between the strands, without significant variation along the sequence. This is observed in the lower panel of Figure 2 that shows a representative configuration of the dodecamer, the water molecules located between the strands, and the ions in the first solvation shell of the duplex. This suggests that an alternative pathway to reproduce the atomistic modulation of the groove width is to tune the water−base and water−sugar interactions. The results presented in subsection B below suggest that the representation of the sugars and, even more, the bases as single interaction sites is too coarse to adequately reproduce the details of their solvation structure.

The thermal stability of DNA was one of the main properties the 3SPN.0 model set out to reproduce. Through a sophisticated parallel tempering procedure, the 3SPN.0 model was parametrized to reproduce experimental melting temperatures of specific sequences. Although we did not include this as part of our parametrization of the mW/3SPN-DNA model, we have performed rudimentary tests to assess the model's stability as well as its ability to be denatured at higher temperatures. The experimental melting temperature of $d(CGCGAATTCGCG)_2$ is approximately 72 °C in 1 M NaCl.[54] In simulations at 85 °C, the explicitly solvated CG DNA distorts significantly by bending and untwisting the duplex; nevertheless, it does not dehybridize (see snapshots in panel b of Figure 3 and rmsd in Figure S8 of Supporting Information). To assess whether this lack of dehybridization is due to the presence of explicit solvent or insufficient sampling, we performed equivalent simulations of the implicitly solvated coarse-grained DNA sequence. A simulation of the 3SPN.0 sequence at 85 °C leads to unzipping within 10 ns. To assess the temperature at which the solvated coarse-grained DNA melts, we performed additional 100 ns simulations at 120, 150, 200, and 250 °C. Denaturation is not observed until

TABLE 2: Radial Distribution $g(r)$ Data for the 1 M NaCl Atomistic (AA) and mW/3SPN-DNA Models

	r (Å) of first shell peak	$g(r)$ value at peak	r (Å) of extent of first shell	first shell neighbors
		O P		
AA	3.6	2.1	4.3	7.4
mW/3SPN-DNA	3.5	1.6	4.2	7.4
		Na P		
AA	3.5	4.6	4.1	0.32
mW/3SPN-DNA	3.4	3.6	4.2	0.32
		Cl P		
AA	6.0	0.5	6.7	0.13
mW/3SPN-DNA	6.3	0.7	7.4	0.25
		O Sugar		
AA	4.6	0.9	6.2	16.5
mW/3SPN-DNA	4.3	1.0	5.9	15.4
		Na Sugar		
AA	6.0	1.8	7.6	1.6
mW/3SPN-DNA	5.6	1.8	7.5	1.2

200 °C (panel c of Figure 3). Interestingly, denaturation is reversible at that temperature: the two strands are caged by the solvent and recombine within the simulation time scale. The coarse-grained solvated dodecamer irreversibly denatures at 250 °C. These results suggest that the melting temperature of the solvated model is between 200 and 250 °C, well above the melting temperature of the implicitly solvated duplex and the experimental value. The dehybridization of the solvated coarse-grained dodecamer starts in the middle region of the sequence. The average distances between complementary bases along the sequence at 25, 150, and 200 °C (presented in Figure S9 of the Supporting Information) show that while the CG ends remain at almost native distances up to 150 °C, the base pairing in the middle region weakens and extends considerably even before denaturation is observed. These results are consistent with the known ability of CG sequences to decrease fraying.

To determine whether the high melting temperature was due to direct (base−base) or indirect (water-induced) interactions, we performed an additional 100 ns simulation of the solvated model at 120 °C in which the base−base interactions between different strands were turned off, and no unzipping was observed. Our results suggest that the coarse-grained bases are too hydrophobic. The explicit inclusion of a water model that likes to associate with itself apparently disfavors the solvation of the hydrophobic base moieties. In principle, this could be corrected by an increase in the water−base attraction. In practice, such an increase in water−base attraction leads to a higher number of water molecules in the hydration shell of the bases that are already overestimated, as will be discussed in section B below. We conclude that the resolution of the bases should be increased, probably to no less than three coarse-grained particles per base, to better match the resolution of the solvent and allow for more favorable solvation free energies without an overall increase in the hydration number of the bases.

B. Solvation Structure. To quantify structural agreement in the solvation structure of the atomistic and coarse-grained DNA, we compared the radial distribution functions (rdfs) of atomistic solutions with those of mW/3SPN-DNA at 1 M NaCl concentration and focused primarily on the location of the first solvation shells of the CG sites. Additionally accessible from these rdfs is the number of neighbors in the first solvation shells, a measure that directly describes the solvation environment of the CG DNA molecule. Table 2 presents the number of first shell neighbors and other signatures from the rdfs for comparison of the

solvation structures of the AA and mW/3SPN-DNA simulations. The rdfs are presented in the Supporting Information (Figures S3−S7).

Hydration Structure. The atomistic simulations show water to have strong affinities to the phosphate group and some of the bases. The first shell surrounding the phosphate moiety contains an average of 7.4 water molecules up to 4.3 Å. Water associates neither closely nor strongly with the sugar groups of the atomistic DNA, giving rise to an extended first shell of water molecules surrounding sugars that are also closely associated with their bridging phosphate groups. An average of 16.5 water molecules can be found within 6.2 Å of the sugar moieties. Water has an interesting interaction with the bases in the atomistic simulations. Thymine, featuring, as a dione, two highly electrophillic sites, shows a very distinguished first shell (see rdfs, Figures S3−S7) extended to 5.2 Å. The other bases show peaks in the rdfs at around 5 Å, yet they do not differentiate a solvation shell as rigorously.

Solvation of the DNA duplex by water is very similar in the mW/3SPN-DNA model and in the AA simulations (Table 2). The interaction between mW water and the CG phosphate is equivalent in terms of the solvating first-shell neighbors: 7.4 at 4.2 Å for the CG, compared with 7.4 at 4.3 Å in AA. The position of the maxima and minima in the first peak of the water−phosphate rdf of the CG and AA models are identical within 0.1 Å. The agreement is almost equally good for the sugar−water interactions: the number of waters in the first shell is 15.4 for the CG and 16.5 for the AA, with a shift in minima and maxima within 0.3 Å. The slight deviations observed in the positions of the first peaks of the water−phosphate and water−sugar rdfs are a result of the coarseness of the model in representing the irregular hydrophilic environment of the atomistic moieties with a single bead. As expected, the more complex the atomistic group represented by a single particle is, the more difficult it is to reproduce its solvation rdf. Thus, it is not surprising that the agreement degrades for the solvation of the bases, represented in the coarse-grained model by a single particle. The mW/3SPN-DNA model exhibits a more featureless rdf of water with the bases compared to the atomistic data (see Supporting Information). For the three bases that do not present well-defined first shell solvation peaks (A, C, and G), the mW/3SPN-DNA model reproduces the water−base distances and the shape of the rdf but systematically overestimates the hydration shell: at 5.75 Å there are about eight water molecules in the hydration shell of all CG bases, while there are about six for the atomistic A, G, and C and ∼4 for T. There is a clear inverse correlation between the complexity of the actual solvation shell of the bases at the atomistic level and the ability of the coarse-grained model to reproduce it with a single site base. When the degree of coarsening of the DNA moiety is comparable to that of water in mW (e.g., phosphate), the agreement in solvation of the DNA by the coarse-grained model is remarkable. Considering the lack of hydrogens and electrostatics in the CG model of water and ions, this approach holds promise for expansion of explicit solvation by mW to other CG biomolecules and suggests that the best results will be obtained when the degree of coarsening of the biomolecule is comparable to the one of the mW model (i.e., beads effectively representing a few atoms).

Ionic Condensation. The condensation of sodium ions around the negatively charged DNA has been extensively studied,[19,24,26−29,55] and is witnessed in the atomistic simulations presented here. In spite of its lack of electrostatics, the coarse-grained model faithfully reproduces the position of the maximum, minimum, and number (0.32) of sodium neighbors in the first shell of the DNA's phosphate groups (Table 2). The sodium ions show very little affinity toward the sugar groups, exhibiting a very extended, featureless shell up to 7.5 Å with an average of 1.6 Na^+ neighbors in the AA and 1.2 in the mW/3SPN-DNA models. The ratio of water to ions in the first solvation shell of the sugars is about 10:1.

The bases T and G show a well-defined peak around 4.5 Å in the atomistic sodium−base rdfs, due to direct contact of the cation with the base. The high degree of coarsening for the bases, again, poses a challenge for an accurate representation of the cation−base rdf in the mW/3SPN-DNA model as the irregular, anisotropic, nucleophilic environment of each base is not sufficiently well-represented by a spherical potential. The mW/3SPN-DNA model cannot capture the Na−base contact peak well: although there is some interaction of Na with the bases at close range—particularly with G—the features of the sodium solvation of the bases are broader and not as defined as the atomistic data (see rdfs in Supporting Information). The success of the mW-ion in replicating the hydration structure of Na and Cl ions suggests that the lack of agreement in mW/3SPN-DNA is mainly due to a too coarse level of description of the bases. The implementation of explicit coarse-grained solvation to more detailed coarse-grained models of DNA such as, for example, that presented in ref 13 or 56 should improve the agreement.

Replicating the different affinities of sodium and chloride ions for DNA is a particular challenge due to the lack of electrostatic interactions in the mW-ion model. Our results for the 1 M solution show that the solvation shell around the phosphate ions is very well reproduced by the coarse-grained model. mW/3SPN-DNA reproduces the number of water, sodium, and chlorides around the phosphates (P) within 7% of the atomistic model. The average distances between the phosphates and the solvent moieties reflect a strong preference of P for water, which on average is even closer to the phosphate than the sodium counterion. The reason is probably a balance of Na−P direct attraction and the electrostatic repulsion that P experiences with Cl that, in turn, associates with Na (this is more readily observed in the cylindrical distribution of ions around DNA, presented below). The result is that the chloride ions are farther from DNA, due to the electrostatic repulsion of the negative charges. This is evident in the radial distribution Cl−P, where the peak moves toward larger distances in the atomistic simulations. This is well-captured by the mW/3SPN-DNA model.

Cylindrical Distributions of Ions around DNA. To further understand the solvation structure of water and ions around the DNA molecule, we calculated cylindrical distributions around the curved central axis of the DNA base pairs. The cylindrical distributions from the atomistic and mW/3SPN-DNA simulations with 1 M NaCl are presented in Figure 4. The mW/3SPN-DNA simulations replicate the features seen in the atomistic model, despite the lack of electrostatics. Measuring across the helix, the width of the coarse-grained DNA molecule is about 1 Å wider than that of the atomistic DNA (on the order of a C−C single bond). This is consistent with the slightly elongated end-to-end distance observed with the CG model. Atomistic data shows water to be closely hydrating the helix, with peaks corresponding to solvating base, sugar, and phosphate moieties at 5, 8, and 12 Å, respectively. The positions of these peaks and their intensities are well-reproduced by the mW/3SPN-DNA model.

The interaction of sodium is very similar between the atomistic and coarse-grained models. The mW/3SPN-DNA

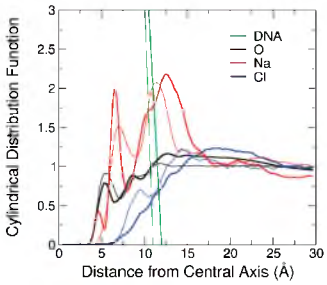


Figure 4. Cylindrical distribution of water and ions around the d(CGCGAATTCGCG) duplex, calculated for each segment along the curvilinear path of the center of sequential base-pairs (i.e., the central axis) and averaged over representative 100+ ns trajectories. Atomistic data are symbolized by thick lines and mW/3SPN-DNA by thin lines. All the DNA CG sites are represented in a single curve (green), with the atomistic showing a wider radius across the helix. Distribution of water (O) is shown in black, sodium (Na) in red, and chloride (Cl) in blue.

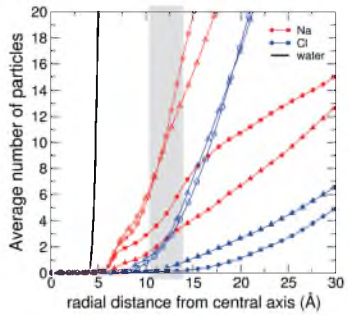


Figure 5. Average number of water, sodium, and chloride atoms surrounding DNA as a function of radial distance from its central axis for 1 M and 200 mM NaCl concentrations. Water, sodium, and chloride correspond to black, red, and blue curves, respectively. Atomistic data are circles and mW/3SPN-DNA triangles. Filled symbols correspond to 200 mM NaCl and empty symbols to 1 M NaCl. Note: the four data sets for water are indistinguishable on the scale of this graph and are therefore presented as a single curve for clarity. The gray box indicates the average location of the phosphate moieties of the DNA strands.

model does not quantitatively reproduce the Na−base distribution (see discussion above), yet the Na profile around DNA is dominated by the interaction with the phosphate and the sugar, which are both well-reproduced by the coarse-grained model (Table 2). Sodium solvation of the bases, sugar, and phosphate moieties appear as three peaks in the cylindrical distribution. A small inner peak at ~4 Å, corresponding to the solvation of the bases by sodium, is not reproduced by the coarse-grained model. This is consistent with the poor representation of the sodium−base rdf discussed above. A second peak around 6 Å is due to the Na−sugar attraction and is well-reproduced by the mW/3SPN-DNA model (the atomistic peak is sharper, but it integrates to the same number of neighbors). The third and most pronounced peak at ~12.5 Å is due to their interaction of sodium with the phosphates. The mW/3SPN-DNA model reproduces the distribution of sodium around DNA of the fully atomistic model in 1 M NaCl. A noteworthy feature of the sodium distribution is how short-ranged it is: the density of ions reaches the average value already within 5 Å of the DNA perimeter. This evidences a high degree of counterion condensation around the DNA, similar to previous studies reported in the literature.[19,24,26−28]

Chloride ions are repelled by the negatively charged atomistic DNA. This is evident in the cylindrical distribution by a depleted chloride density (i.e., the distribution function has a value less than 1) within the region occupied by DNA. On the other hand, the chloride anions are attracted to the sodium cations, leading to a layering of charged species of which the chloride forms the outer shell. This produces a modest increase in the density of Cl^- outside the outermost Na^+ peak (Figure 4). The same features (depletion within the DNA region and slight increase outside the sodium peak) are observed in the cylindrical distribution of chloride in the mW/3SPN-DNA model. The small peak around 8 Å in the depleted region of the coarse-grained model is due to a stronger Na−Cl attraction in mW-ion model than in the atomistic model. The Na−Cl interaction was purposefully decreased in the atomistic model[57] to avoid crystallization of the salt. The mW-ion model, which has a stronger Na−Cl interaction than the atomistic one, reproduces the experimental[58] association constant of NaCl in water.[31]

The agreement between atomistic and coarse-grained models is the best in concentrated electrolyte solutions. This is appreciated in Figure 5, which presents the integrated average number of solvent molecules surrounding the central axis of the DNA strand as a function of the distance to the helical axis. The agreement between the atomistic and mW/3SPN-DNA simulations is excellent for the 1 M solution, as seen also in Figure 4. Simulations at 0.2 M salt concentration show, both in the atomistic and coarse-grained models, qualitatively similar trends when compared to the 1 M solutions: there is less sodium condensation around the DNA in the dilute solution, and the chlorides are farther from the duplex. The change in ionic condensation from 0.2 to 1 M is not in accordance with what is expected by counterion condensation theory (valid at low electrolyte content),[24,25] which predicts the neutralization of the polyelectrolyte chain to be mostly invariant with salt concentration. Although the coarse-grained model reproduces the qualitative distribution of NaCl around the double helix for the 0.2 M solution, the agreement with the atomistic model is not quantitative as for the 1 M one: Figure 5 shows that at 0.2 M the CG model under-represents the amount of Na ions condensed onto the DNA and over-represents the distribution of chloride close to it. This degradation of the agreement in the more dilute solution is a limitation of the use of short-ranged potentials for the description of the interactions between ions.

The comparison between the water and ion distribution around the atomistic and mW/3SPN-DNA data is very favorable, especially for solvation by 1 M NaCl. This may be surprising, considering that DNA is a polyelectrolyte, the model has no charges or long-ranged interactions, and that this distribution is not a feature that we included in refining the parameters of the model. The good agreement in the concentrated solution is due to the lack of long-range ionic structure around the DNA, as the atomistic ions condense around the highly charged polyelectrolyte. When the polyelectrolyte is surrounded by water and condensed cations, the effective length of the interactions, described in the atomistic regime by long-ranged Coulombic interactions, is reduced. When the ionic concentration decreases, on the other hand, and the ionic distribution extends farther from the polyelectrolyte, the coarse-grained model still captures the qualitative distribution of ions but underestimates the strong depletion of Cl around the helix because being short-ranged it cannot account correctly for the less shielded, longer ranged,

TABLE 3: Residence Time (in ps) for the 1 M NaCl Atomistic (AA) and mW/3SPN-DNA Models

moieties	AA	mW/3SPN-DNA[a]
O−P	38	33
Na−P	210	100
Cl−P	26	33
O−sugar	43	43

[a] CG residence times were multipled by 2.5 to compare to AA data (see the text).

repulsion between the phosphates and the chloride. The mW-ion model that describes the solvent in this work shows good transferability from 0.66 to 5.6 M NaCl, the range of its original parametrization. It may be possible to parametrize the interactions to reproduce more faithfully the distribution of ions for dilute solutions; we expect that this would result in longer ranged interactions that would compromise the efficiency of the model.

C. Solvation Dynamics. We used the residence time of water and ions around the DNA molecule to determine how well the mW/3SPN-DNA model captures the solvation dynamics. This is an important feature of the solvation, as it may control the accessibility of other molecules (e.g., drug molecules, proteins, other nucleic acids) to DNA. Comparing the relative times that the water and ions are associated with the DNA moieties in the atomistic model gave us also a measure of the relative interaction strengths for use in developing the solvent−DNA parameters. These calculations are presented in Table 3 for the various interactions of water, ions, and the DNA backbone in the 1 M NaCl solution. In comparing atomistic data with that of the mW/3SPN-DNA model, we point out that CG models are inherently faster than their atomistic counterparts. This is due to the evolution of the CG system on a smoother potential energy surface. For the mW-ion model,[31] we found the diffusion of water in 1 M NaCl at 300 K to be 2.5 times faster than the atomistic (TIP4P-Ew). Thus, using the diffusivity of water to normalize the time scales, our target was to obtain residence times in the mW/3SPN-DNA model that are 2.5 times shorter than in the fully atomistic simulation. We therefore multiply the residence times we calculated for the mW/3SPN-DNA model by this factor so as to compare the dynamics of the two models directly. The residence times were computed, following the protocols of ref 31, as the average time it takes for the ion (water) to move to a distance further than the first minimum of the rdf of the DNA−ion (DNA−water). We computed residence times only for pairs that form well-defined first peaks in the rdf.

The shortest residence times are, not unexpectedly, between the anions: the residence time of Cl around phosphate is ∼30 ps. This is comparable to ∼35 ps, the average time it takes for water to diffuse over 7 Å, the width of the first peak of the Cl−P rdf. We note that this first peak encompasses distances for which Na or water mediates the two anions; we do not find contact pairs between chloride and phosphates.

The residence times for water on the phosphate and sugar are similar at around 40 ps. As the width of the water−phosphate peak in the rdf is about 4.2 Å, this residence time is about twice as long as the diffusional time for water in the bulk over the same distance, consistent with a strong water−phosphate attraction. The width of the water−sugar peak is about 7 Å; thus, these results indicate that water is not particularly attached to the sugars of DNA.

The longest residence times are for Na with the phosphate ions of the DNA, as is expected due to their strong attraction. The coarse-grained model underestimates the residence time for Na−P. Although mW/3SPN-DNA reproduces the number of neighbors and location of maxima and minima of the first peak, it predicts the peak to be less sharp than for the atomistic model. It should be explored in the future whether the number of neighbors and the scaled residence times for Na−P can be simultaneously reproduced by the coarse-grained model. Except for the Na−P pair, the scaled residence times of the mW/3SPN-DNA are within 10−20% of the atomistic; thus, we expect that processes dominated by desolvation of DNA should be qualitatively reproduced by the coarse-grained model.

D. Benchmarking. Comparison of mW/3SPN-DNA to all-atom simulations shows a sizable increase in computational performance. For the system sizes of this study, the explicitly solvated coarse-grained model is 90 times faster than the atomistic model (i.e., it takes 90 times more CPU time to compute a given simulation time, e.g., 1 ns, with the atomistic model). Most of the expense of the mW/3SPN-DNA simulations is from the explicit representation of water and ions, which, as we showed for mW-ion,[31] scales approximately as $N \log N$ (where N is the number of coarse-grained particles in the system). It is possible for our simulations with mW/3SPN-DNA to extend into the microsecond time frame with only 7 days of wall time across 48 Intel Xeon E5430 (2.66 GHz) processors for a double-stranded DNA solvated by ca. 16 000 molecules of water and ions.

4. Conclusions

In this work, we have combined a coarse-grained model of DNA with a coarse-grained model of water and ions to produce a coarse-grained model of DNA with explicit solvation. Using an established CG model of DNA—the 3SPN.0 model of Knotts et al.—allowed us to tap the wealth of knowledge already carefully developed into its force field and parameters. Our focus was mostly involved in adapting the mW-ion model—developed to introduce coarse-grained ions to the monatomic water (mW) model—to interact with the DNA molecule so as to retain its stability yet also exhibit the nuances of solvation. The multitude of parameters of the mW/3SPN-DNA model has not been fully explored, yet the results show a very good empirical match of the solvation structure to the atomistic reference. The model and parameters presented are our best empirical fit to the atomistic data.

Replicating the structure of the solution of ions and water surrounding the DNA molecule was a primary goal of parametrization and is measurable from the radial distribution functions with features such as the location of the solvation shells and number of solvating neighbors. Our experience with the mW-ion model allowed us to describe well the solvation of the phosphates with the coarse-grained water and ions, despite the use of extremely short-ranged forces (i.e., no electrostatics). The results suggest that the mW/3SPN-DNA model is able to capture the long-ranged structure of water and ions around the DNA molecule, as shown by the cylindrical distribution. Moreover, the explicit solvation of the CG DNA actually helps stabilize the molecule, as evidenced by the damping of its structural fluctuations and the increase in the melting temperature.

We found an inverse correlation between the ability of the coarse-grained model to describe the radial distribution functions and the degree of coarsening of the DNA moiety involved; the agreement between atomistic and mW/3SPN-DNA models was very good for rdfs involving phosphates, slightly less for the sugars, and the least for the bases. Despite the differences in size and complexity of the atomistic representations of phosphate, sugar, and the four bases, each of these is described by

a single particle in the present CG model. The very good agreement in the description of the water−water, water−salt, and water−phosphate structures indicates that the model could be improved by increasing the resolution of the bases and the sugars. An increase in the number of particles per coarse-grained base and sugar (to, for example, two and three for the bases, as in ref 56 or three for sugar as in the M3B model[59]) would marginally affect the efficiency of the simulations, as most of the computational cost is attributable to the solvent.

Ions and solvent play a crucial role in the melting transition: Schatz and co-workers have shown that the presence of explicit ions in coarse-grained simulations of DNA sharpens the melting transitions,[18,60] and De Pablo and co-workers have shown that addition of a mid-range sugar−sugar attractive term to the implicitly solvated coarse-grained DNA force field achieves quantitative melting temperatures and a good description of the renaturation process.[38−40] It may thus seem paradoxical that a model with explicit solvation, such as mW-DNA/3SPN, performs less well in the prediction of the melting temperatures than the implicitly solvated 3SPN. Our results strongly suggest that the high melting temperature of the mW/3SPN-DNA Dickerson dodecamer (between 200 and 250 °C, higher than for the implicitly solvated 3SPN model and the experiment) is due to a strong solvent-induced hydrophobic attraction between the bases in the presence of explicit water. An increase in the water−base attraction would lead to a lower melting point, by favoring the solvation of the single strands. Nevertheless, we find that the current parameters already overestimate the number of water molecules in the hydration shell of the bases, and this would be exacerbated with an increase in water−base attraction. This poses a representability problem: the model cannot simultaneously represent the melting temperature and the solvation of the bases. We conjecture that commensurate accurate melting temperatures and solvation could be obtained through a finer resolution of the bases (at least 3 sites per base) that would allow for the distinction of its internal hydrophilic and hydrophobic moieties and a better resolution of their shape (i.e., excluded volume) as experienced by a water molecule.

The results of this work show that a coarse-grained model with very short ranged interactions, no hydrogen atoms, and no electrostatics is able to reproduce rather quantitatively the distribution and relative residence times of water and ions around DNA. The strategy developed in this work could be applied to the coarse-grained solvation of proteins, RNA, and other biomolecules. We expect the solvation of proteins to be the most challenging, because of the fact that the density of ionic moieties is generally much lower in proteins than in DNA and that part of the success of the present model is based on the condensation of the counterions around the polyelectrolyte. We envision the explicitly solvated mW/3SPN-DNA model to have the greatest applicability in model systems where hydrophobic/hydrophilic interactions and solvation play a large role in the microscopic behavior and the role of the solvent cannot be ignored. Examples of this include interactions of DNA with drug molecules, surfaces, micelles, and proteins.

Acknowledgment. We thank Thomas A. Knotts IV and Terry Schmitt for sharing the 3SPN code and for insightful comments and discussions. We acknowledge the support of the National Science Foundation through Collaborative Research Grant CHE-0628257 (VM), a TeraGrid award of computer time TG-MCA01S027 (TEC3), and the National Institutes of Health GM-081411 (TEC3). We thank the Center of High Performance Computing at the University of Utah for allocation of computing time.

Supporting Information Available: Expanded data sets characterizing the stability of the DNA molecule, graphed radial distributions, the full parameter set of the mW/3SPN-DNA model, and analysis of simulations of DNA solvated in a 200 mM NaCl solutions. This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III *Acc. Chem. Res.* **2000**, *33*, 889.
(2) Cheatham, T. E., III *Curr. Opin. Struct. Biol.* **2004**, *14*, 360.
(3) Orozco, M.; Noy, A.; Perez, A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 185.
(4) Mackerell, A. D., Jr.; Nilsson, L. *Curr. Opin. Struct. Biol.* **2008**, *18*, 194.
(5) Hashem, Y.; Auffinger, P. *Methods* **2009**, *47*, 187.
(6) Beveridge, D. L.; McConnell, K. J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 182.
(7) Kenward, M.; Dorfman, K. D. *J. Chem. Phys.* **2009**, *130*, 10.
(8) Malikova, N.; Cadene, A.; Marry, V.; Dubois, E.; Turq, P.; Zanotti, J. M.; Longeville, S. *Chem. Phys.* **2005**, *317*, 226.
(9) Bruant, N.; Flatters, D.; Lavery, R.; Genest, D. *Biophys. J.* **1999**, *77*, 2366.
(10) Locker, R.; Harvey, S. C. *Biophys. J.* **2005**, *88*, 229A.
(11) Olson, W. K.; Zhurkin, V. B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 286.
(12) Trovato, F.; Tozzini, V. *J. Phys. Chem. B* **2008**, *112*, 13197.
(13) Poulain, P.; Saladin, A.; Hartmann, B.; Prevost, C. *J. Comput. Chem.* **2008**, *29*, 2582.
(14) Morriss-Andrews, A.; Rottler, J.; Plotkin, S. S. *J. Chem. Phys.* **2010**, *132*, 035105.
(15) Knotts, T. A.; Rathore, N.; Schwartz, D. C.; de Pablo, J. J. *J. Chem. Phys.* **2007**, *126*, 084901.
(16) Drukker, K.; Schatz, G. C. *J. Phys. Chem. B* **2000**, *104*, 6108.
(17) Knotts, T. A.; Rathore, N.; Schwartz, D. C.; de Pablo, J. J. *J. Chem. Phys.* **2007**, *126*, 12.
(18) Lee, O.-S.; Prytkova, T. R.; Schatz, G. C. *J. Phys. Chem. Lett.* **2010**, *1*, 1781.
(19) Cheatham III, T. E.; Young, M. A. *Biopolymers* **2000**, *56*, 232.
(20) Stefl, R.; Cheatham, T. E.; Spackova, N.; Fadrna, E.; Berger, I.; Koca, J.; Sponer, J. *Biophys. J.* **2003**, *85*, 1787.
(21) Spackova, N.; Cheatham, T. E.; Ryjacek, F.; Lankas, F.; van Meervelt, L.; Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2003**, *125*, 1759.
(22) Basdevant, N.; Ha-Duong, T.; Borgis, D. *J. Chem. Theory Comput.* **2006**, *2*, 1646.
(23) Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2005**, *122*, 124906.
(24) Manning, G. S. *J. Chem. Phys.* **1969**, *51*, 924.
(25) Manning, G. S. *Q. Rev. Biophys.* **1978**, *11*, 179.
(26) Spasic, A.; Mohanty, U. Counterion condensation in nucleic acid. In *Advances in Chemical Physics*; John Wiley & Sons Inc: New York, 2008; Vol. 139; pp 139.
(27) Volker, J.; Klump, H. H.; Manning, G. S.; Breslauer, K. J. *J. Mol. Biol.* **2001**, *310*, 1011.
(28) Muthukumar, M. *J. Chem. Phys.* **2004**, *120*, 9343.
(29) Record, M. T.; Zhang, W. T.; Anderson, C. F. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. In *Advances in Protein Chemistry*; Academic Press Inc: San Diego, 1998; Vol. 51; pp 281.
(30) Hud, N.; Polak, M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 293.
(31) DeMille, R. C.; Molinero, V. *J. Chem. Phys.* **2009**, *131*, 034107.
(32) Molinero, V.; Moore, E. *J. Phys. Chem. B* **2009**, *113*, 4008.
(33) Moore, E. B.; Molinero, V. *J. Chem. Phys.* **2009**, *130*, 244505.
(34) Jacobson, L. C.; Hujo, W.; Molinero, V. *J. Phys. Chem. B* **2009**, *113*, 10298.
(35) Jacobson, L. C.; Molinero, V. *J. Phys. Chem. B* **2010**, *114*, 7302.
(36) Moore, E. B.; Molinero, V. *J. Chem. Phys.* **2010**, *132*, 244504.
(37) Jacobson, L. C.; Hujo, W.; Molinero, V. *J. Am. Chem. Soc.* **2010**, *132*, 11806.
(38) Sambriski, E. J.; Schwartz, D. C.; de Pablo, J. J. *Biophys. J.* **2009**, *96*, 1675.
(39) Sambriski, E. J.; Ortiz, V.; de Pablo, J. J. *J. Phys.: Condens. Matter* **2009**, *21*, 034105.
(40) Sambriski, E. J.; Schwartz, D. C.; de Pablo, J. J. *Proc. Natl. Acad. Sci.* **2009**, *106*, 18125.
(41) Arnott, P. J. S.; Smith, C.; Chandrasekaran, R. Atomic coordinates and molecular conformations for DNA−DNA, RNA−RNA, and DNA−

RNA helices. In *CRC Handbook of Biochemistry and Molecular Biology*: 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, 1976; Vol. II; pp 411.

(42) Shui, X.; McFail-Isom, L.; Hu, G.; Williams, L. *Biochemistry* **1998**, *37*, 8341.

(43) Stillinger, F. H.; Weber, T. A. *Phys. Rev. B* **1985**, *31*, 5262.

(44) Carre, A.; Berthier, L.; Horbach, J.; Ispas, S.; Kob, W. *J. Chem. Phys.* **2007**, *127*, 114512.

(45) Tjandra, N.; Tate, S.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190.

(46) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick. T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.

(47) Joung, I. S.; Cheatham, T. E., III *J. Phys. Chem. B* **2008**, *112*, 9020.

(48) Plimpton, S. J. *J. Comput. Phys.* **1995**, *117*, 1.

(49) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California, San Francisco, 2008.

(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Comput. Phys.* **1984**, *81*, 3684.

(51) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 4.

(52) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(53) Tjandra, N.; Tate, S.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190.

(54) Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Broka, C.; Rice, J. A.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 428.

(55) Dixit, S. B.; Beveridge, D. L.; Case, D. A.; Cheatham, T. E.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; Thayer, K. M.; Varnai, P. *Biophys. J.* **2005**, *89*, 3721.

(56) Dans, P. D.; Zeida, A.; Machado, M. R.; Pantano, S. *J. Chem. Theory Comput.* **2010**, *6*, 14.

(57) Joung, I. S.; Cheatham, T. E., III *J. Phys. Chem. B* **2008**, *112*, 9020.

(58) Fuoss, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 34.

(59) Molinero, V.; Goddard, W. A. *J. Phys. Chem. B* **2004**, *108*, 1414.

(60) Prytkova, T.; Eryazici, I.; Stepp, B.; Nguyen, S.; Schatz, G. C. *J. Phys. Chem. B* **2010**, *114*, 2627.

JP107028N

Supporting Information For

"A Coarse-Grained Model For Explicit Solvation Of Dna By Water And Ions"

Robert C. DeMille, Thomas E. Cheatham, and Valeria Molinero

Table S1. Stillinger-Weber Force Field Parameters of the mW/3SPN-DNA model

Stillinger-Weber **pair** potentials				
i	***j***	ε_{ij} ***(kcal/mol)***	σ_{ij} ***(Å)***	a_{ij}
mW	mW	6.189	2.3925	1.8
mW	Na	16.00	1.85	1.8
mW	Cl	15.00	2.60	1.8
mW	P	4.0	3.05	1.7
mW	S	0.2	3.5	1.8
mW	A	0.5	4.0	1.8
mW	T	0.6	4.1	1.8
mW	C	0.6	4.1	1.8
mW	G	0.4	4.0	1.8
Na	Na	0	3.3	1.1
Na	Cl	10.2	2.00	1.8
Na	P	3.5	2.6	1.8
Na	S	0.2	3.2	1.8
Na	A	0.5	3.5	1.8
Na	T	8.5	3.3	1.6
Na	C	0.5	3.5	1.8
Na	G	10.0	3.3	1.6
Cl	Cl	0	3.9	1.1
Cl	P	0	5	1.0
Cl	S	1.0	6.0	1.1
Cl	A	0.2	5.0	1.8
Cl	T	0.2	4.5	1.8
Cl	C	0.2	4.5	1.8
Cl	G	0.2	4.5	1.8
P	P	0	4.0	1.0
Parameters same for all ij pairs: A=7.049556277, B=0.602224558, p=4, q=0, γ=1.2				

Table S1. Stillinger-Weber Force Field Parameters of the mW/3SPN-DNA model (continuation)

Stillinger-Weber **3-body** potentials (i is center atom, with θ_0 forming the angle between legs r_{ij} and r_{ik}. ε_{ijk} *units kcal/mol*)					
i	***j***	***k***	ε_{ijk}	λ_{ijk}	$cos\theta_{ijk}$
mW	mW	mW	6.189	23.15	-1/3
mW	mW	Na	16	7	-1/3
mW	mW	Cl	15	16	-1/3
mW	mW	P	12	12	-1/3
mW	Na	Na	16	7	-1/3
mW	Na	Cl	18.2	1	-1/3
mW	Na	P	5	1	-1/3
mW	Cl	Cl	15	16	-1/3
mW	Cl	P	0	0	-1/3
mW	P	P	4	0	-1/3
Na	mW	mW	16	7	-1/3
Na	mW	Na	16	7	-1/3
Na	mW	Cl	18.2	1	0.74606
Na	mW	P	5	1	-1/3
Na	Na	Na	0	0	-1/3
Na	Na	Cl	0	0	-1/3
Na	Na	P	0	0	-1/3
Na	Cl	Cl	10.2	0	-1/3
Na	Cl	P	5	1	1
Na	P	P	3.5	0	-1/3
Cl	mW	mW	15	16	-1/3
Cl	mW	Na	18.2	1	0.87648
Cl	mW	Cl	15	16	-1/3
Cl	mW	P	0	0	-1/3
Cl	Na	Na	10.2	0	-1/3
Cl	Na	Cl	0	0	-1/3
Cl	Na	P	0	0	-1/3
Cl	Cl	Cl	0	0	-1/3
Cl	Cl	P	0	0	-1/3
Cl	P	P	0	0	-1/3
P	mW	mW	5.5	0	-1/3
P	mW	Na	5	1	-1/3
P	mW	Cl	0	0	-1/3
P	mW	P	12	12	-1/3
P	Na	Na	3.5	0	-1/3
P	Na	Cl	0	0	-1/3
P	Na	P	0	0	-1/3
P	Cl	Cl	0	0	-1/3
P	Cl	P	0	0	-1/3
P	P	P	0	0	-1/3
3 body parameters for configurations of ijk are the same as ikj.					

Table S2. Yukawa Force Field Parameters of the mW/3SPN-DNA model

i	*j*	*C* (*kcal/mol*)	κ (Å^{-1})	r_{cut} (Å)
Na	Na	1107	1.8	7.0
Cl	Cl	1107	1.8	7.0
Cl	P	2107	1.2	9.0

Table S3. DNA structural stability of the three models used in this study for 200 mM NaCl from representative 100 ns trajectories.

Atomistic	3SPN.0	mW/3SPN-DNA
End-to-End Distance (Å)		
29.7±0.1	33.8±0.2	34.2±0.2
RMSD (Å) from average structure		
1.5±0.4	2.7±0.8	2.6±0.7
RMSD (Å) from standard B-form		
2.4±0.4	4.7±1.0	4.2±0.5

Figure S1. Angle formed between DNA moieties averaged over 100ns+ trajectories from the DNA models used in this study (20 ns for atomistic data). X axis corresponds to the type of angle in the coarse-grain force field, i. e. Table III of the 3SPN.0 model [Knotts, T. A. IV; et al. *J. Chem. Phys.* **2007**, 126, 084901]. The reference structure data (black circles) is based on the standard model of B form dsDNA (see text). CG data (implicit 3SPN.0 model in red squares, explicitly solvated mW/3SPN-DNA in blue triangles) closely aligns with this data, as the model was parameterized to fit it. Atomistic data (purple diamonds) displays drastically different average values for the angles, showing the CG data to overestimate the sugar-phosphate backbone angles (types 1 and 2), and generally underestimate the phosphate-sugar-base angles (types 3-10).

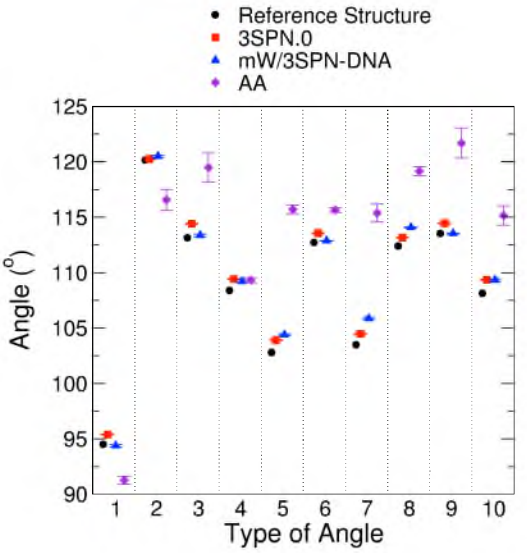
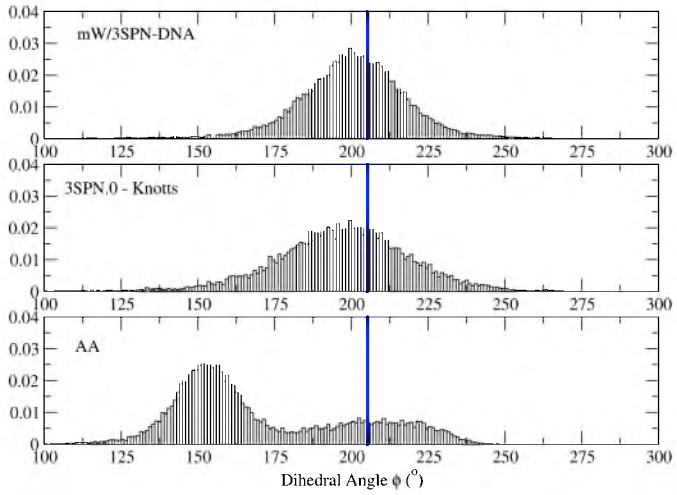
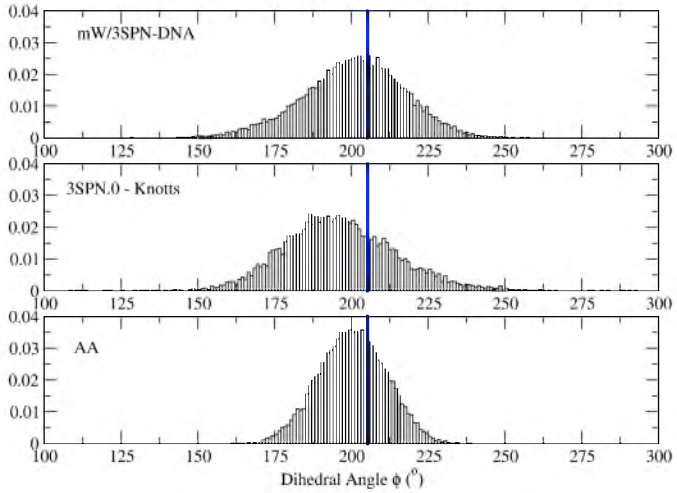


Figure S2. Distribution of dihedral angles formed between DNA moieties along the phosphate-sugar backbone averaged over 100ns+ trajectories from the DNA models used in this study (20 ns for atomistic data). Negative dihedral angles have been shifted by +360° to compare the angular distribution on a continuous scale of 0-360° instead of -180-180°. Blue lines indicate the dihedral angle formed in the standard form of B form dsDNA, which was the target parameter in the CG 3SPN.0 model [Knotts, T. A. IV; et al. *J. Chem. Phys.* **2007**, 126, 084901], corresponding to Table III of that report (shifted by +360°). The implicit 3SPN.0 model and the explicitly solvated mW/3SPN-DNA model show excellent agreement in the dihedral angular distribution, though differ from the atomistic distribution, which sometimes exhibits bimodality (e.g. panels A & C).

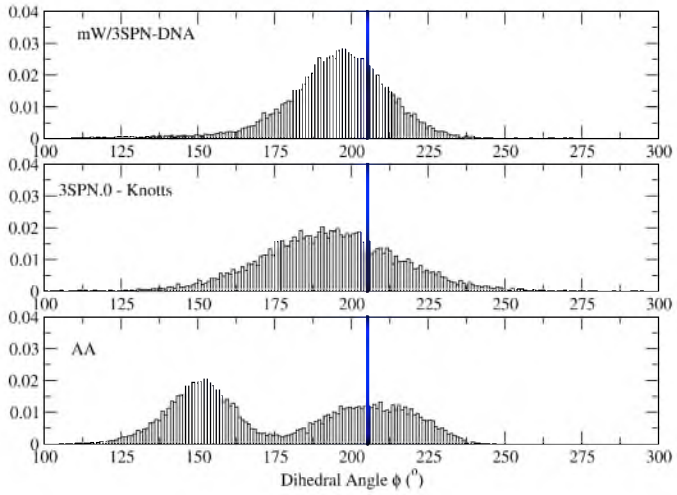
Panel A. Dihedral distribution for P-(5')S(3')-P-(5')S through the first CGCG section of the Dickerson dodecamer.



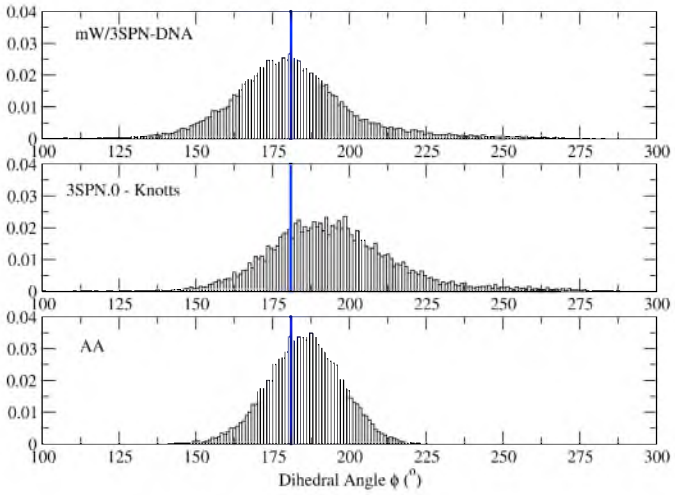
Panel B. Dihedral distribution for P-(5')S(3')-P-(5')S through the middle AATT section of the Dickerson dodecamer.



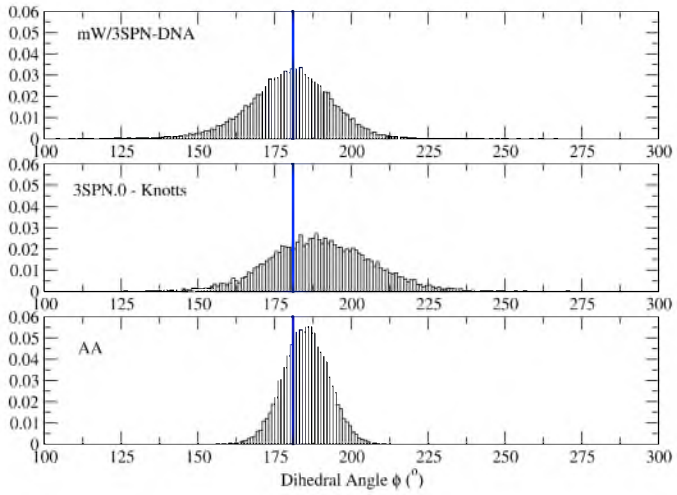
Panel C. Dihedral distribution for P-(5')S(3')-P-(5')S through the second CGCG section of the Dickerson dodecamer.



Panel D. Dihedral distribution for S(3')-P-(5')S(3')-P through the first CGCG section of the Dickerson dodecamer.



Panel E. Dihedral distribution for S(3')-P-(5')S(3')-P through the middle AATT section of the Dickerson dodecamer.



Panel F. Dihedral distribution for S(3')-P-(5')S(3')-P through the middle second CGCG section of the Dickerson dodecamer.

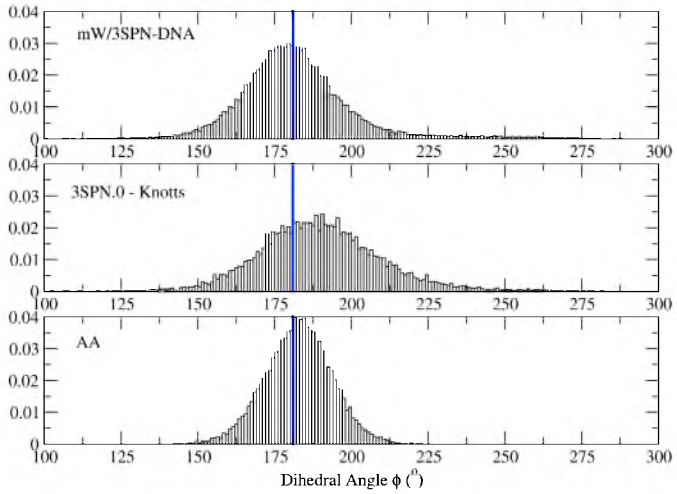


Figure S3. Radial distribution function of water with DNA moieties. Thick solid lines for atomistic data, skinny lines for mW/3SPN-DNA.

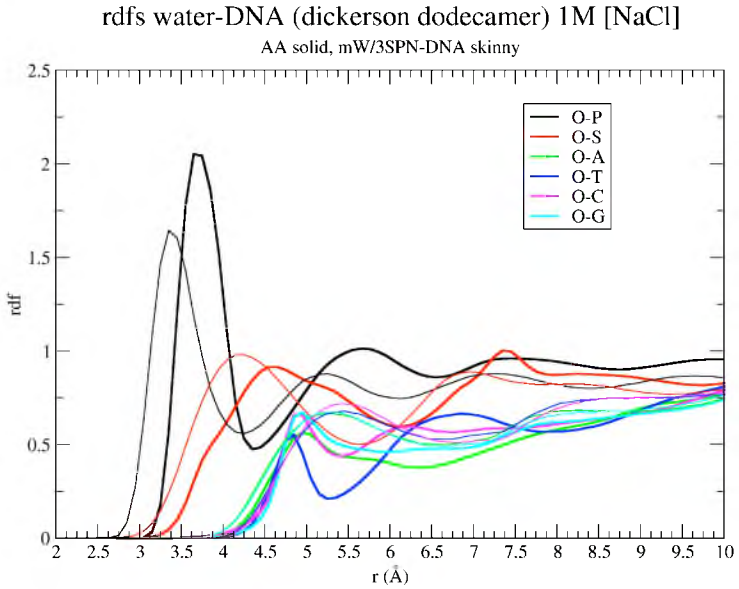


Figure S4. Radial distribution function of sodium with DNA moieties. Thick solid lines for atomistic data, skinny lines for mW/3SPN-DNA.

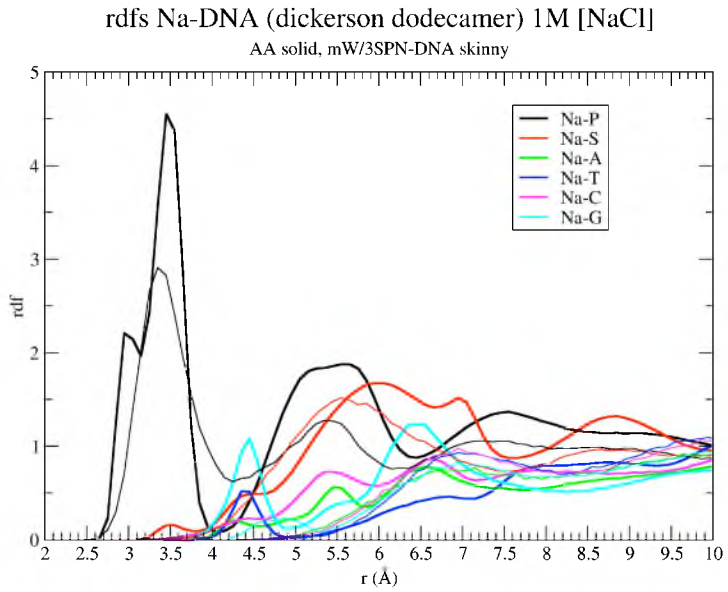


Figure S5. Radial distribution function of chloride with DNA moieties. Thick solid lines for atomistic data, skinny lines for mW/3SPN-DNA.

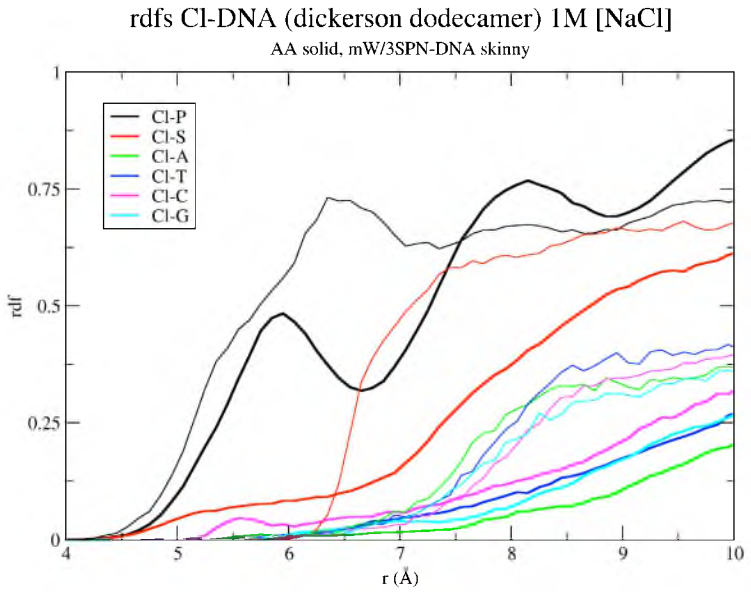


Figure S6. Number of water neighbors around DNA moieties. Thick solid lines for atomistic data, skinny lines for mW/3SPN-DNA.

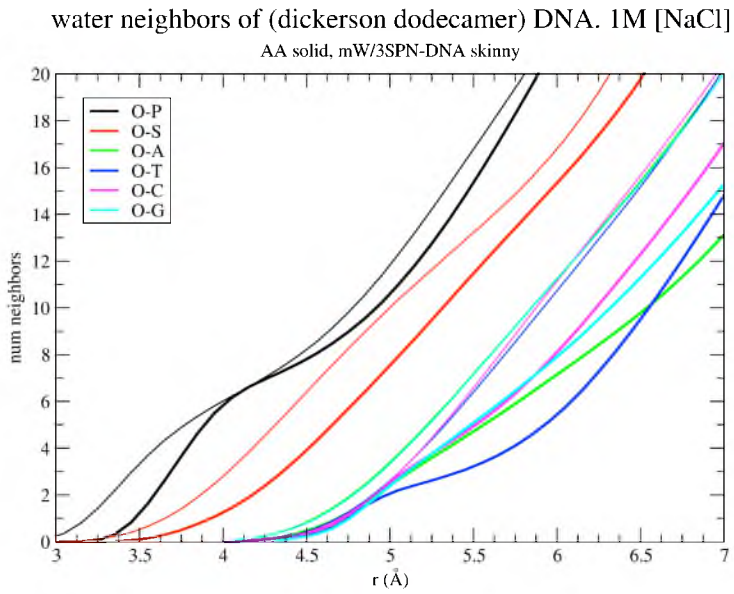


Figure S7. Number of sodium neighbors around DNA moieties. Thick solid lines for atomistic data, skinny lines for mW/3SPN-DNA.

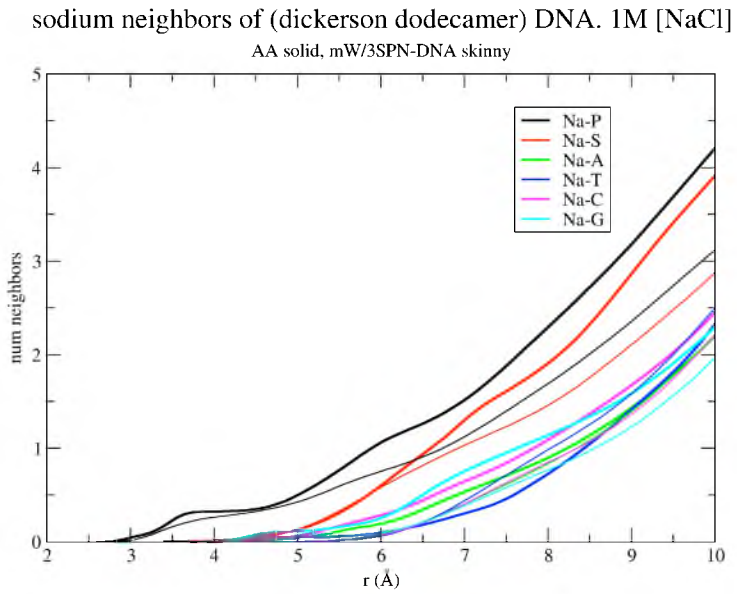


Figure S8. Upper panel: RMSD vs. time for mW/3SPN-DNA simulations of the Dickerson DNA dodecamer in 1M NaCl at 85°C. Lower Panels: Corresponding representative structures at the times labeled in the upper panel as a), b), c) and d). Only minor fluctuations around the equilibrium structure are observed at 25°C

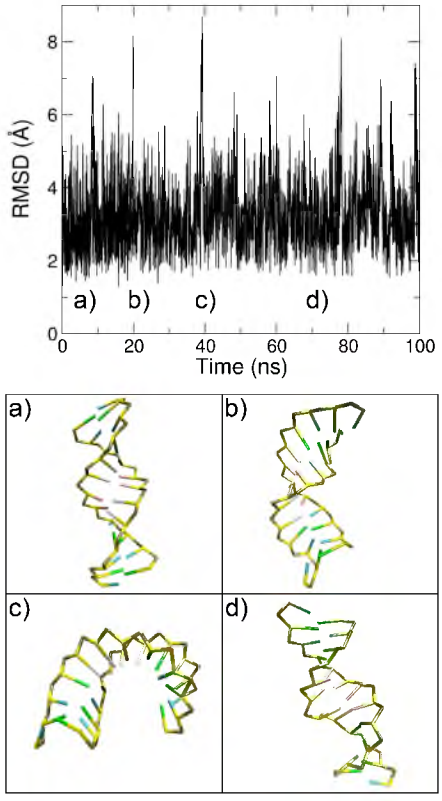
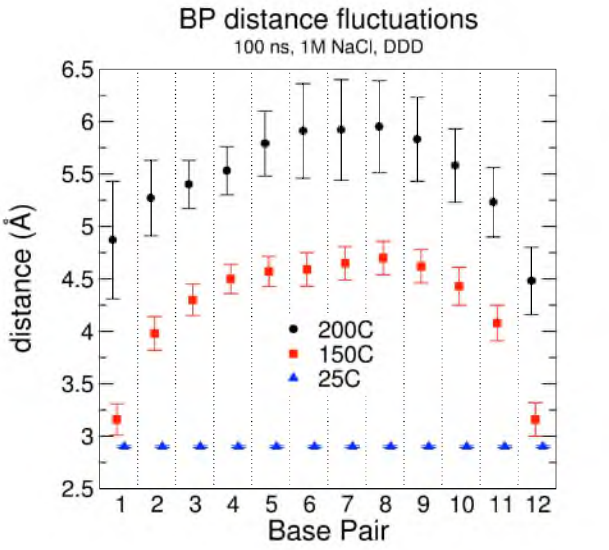


Figure S9. Distance between complementary pairs of bases in the two strands along the dodecamer sequence averaged over 100 ns simulations of the coarse-grained solvated DNA at 25, 150 and 200°C. The CG pairs at the end of the dodecamer prevent fraying, and eventually, at 200 and 250°C dehybridization starts in the middle of the sequence leading to reversible (200°C) or irreversible (250°C) denaturation of DNA.



CHAPTER 4

COMB-BRANCHED POLYMER ELECTROLYTES MODIFIED WITH SIDE-CHAIN SPACERS PROMOTE INCREASED LITHIUM TRANSPORT

Note that this chapter is under preparation for *Macromolecules*.

Introduction

The ideal electrolyte for electronic or automotive purposes should be easy to fabricate, allow rapid diffusion of the charge carriers, and maintain stability across multiple cycles of charging and recharging. Improving lithium batteries has been the focus of many studies [42-45] using solid polymer electrolytes (SPEs), the most prevalent of which is poly(ethylene oxide) (PEO).[46] These experimental and theoretical studies have shown PEO to have disadvantages, especially at lower temperatures, for the ionic conductivity of the SPE systems. To overcome limitations of simple PEO electrolytes, changes to the polymer architecture have been explored[47-54] to organize the SPE in a way to promote conductivity and still retain PEO's desirable properties.

A particular architecture that allows a large array of properties to be explored is a PEO-based comb-branched structure. Comb-branched SPEs have been developed that allow for more customization of physical properties and optimization of ionic conductivity with a wide range of chemical variability.[51-59] To understand the mechanism

of Li^+ transport in these comb-branched systems, theoretical studies[41, 60] have been employed to view the molecular-level workings of the SPE systems. Assuming the model is able to accurately represent the SPE system, using computer simulations can be a relatively inexpensive method to probe potential improvements to SPE architectures.

To investigate how the comb-branched polymer poly(epoxide ether) (PEPE) compared to PEO and to experimental results of PEPE, a simulation study[41] was carried out with the MD package Lucretius[4]. For this study, a force field[61, 62] was used which had been rigorously derived from quantum mechanics for this many-body system and was shown to be accurate for predicting properties of electrolyte systems such as ion transport.[40] With simulation data spanning tens of nanoseconds, this study showed the $PEPE_5$ SPE systems to exhibit lower ionic conductivity and decreased diffusion rates of Li^+ compared to linear PEO. Mechanistically, this was attributable to two differing behaviors seen in the Li^+ cations; some "slow" Li^+ cations were coordinated for long timeframes to the backbone while other "fast" cations hopped quickly between sidechains. Indeed, by eliminating attractive interactions between Li^+ and the backbone ether oxygens, an unphysical but accessible ability of simulation, the cation transport was improved.

In order to continue the study from 2007 involving $PEPE_5$ chains, the present study set out to use the same accurate force field to predict a molecular design that would overcome the limitations found using that particular polymer electrolyte. Adding non-coordinating CH_2 spacers to the attachment points of the backbone, it was hypothesized, would attempt to decouple the motion of the Li^+ cations along the side-chains and the backbone by preventing cation aggregation at these oxygen-rich junctions. The structure

of the spacer-PEPE (sPEPE) polymer studied here is shown in Figure 4.1. A comparison between the PEPE and sPEPE systems results, showing differences in the transport of Li^+, location of Li^+ coordination, as well as chain dynamics and $TFSI^-$ behavior which will be explained in detail below.

Methods

Using the many-body polarization force field[61, 62] developed for the Lucretius[4] simulation package, MD simulations were performed on the LiTFSI-doped, spacer-attached comb-branched poly(epoxide ether) systems. 21 LiTFSI and three sPEPE chains comprised the three-dimensional, periodic cubic simulation cell, and the systems were equilibrated at a pressure of 1 atm for 2-3 ns followed by productions runs of 80 ns for each temperature in the *NVT* ensemble. Details of the temperatures studied and their production run lengths for the sPEPE and original PEPE systems are available in Table 4.1. As with the PEPE systems, each sPEPE chain consists of 20 repeat units, each with 7 ether oxygens, with the dopant concentration at a ratio of 1:20 of Li^+ cations to EOs.

The temperature and pressure of the system were regulated by the Nose-Hoover thermostat and barostat[63] with time constants of 0.1 and 1.0 ps, respectively. Constraint of the bond lengths was managed through the SHAKE algorithm.[64] Long ranged electrostatics involving partial charges and induced dipoles were handled by Ewald summation with an Ewald convergence factor of 8 Å with 6 reciprocal space vectors. The induced dipole – induced dipole interactions were scaled to zero between 9 and 10 Å using a tapering function. A multiple time step reversible reference system propagator algorithm[63] was used; bonding, bending and torsional motions used a time step of 0.5 fs, while nonbonded interactions used a time step of 1.5 and 3.0 fs for interactions less than

6.5 Å and greater than 6.5 but less than 10 Å, respectively. A 3.0 fs time step was also used for the reciprocal space calculations of the Ewald summation.

Results and Discussion

Structural Comparison Between Systems

To understand how the sPEPE system could be instrumental in increasing Li^+ transport, it was first desirable to study the differences in structure prompted by replacing PEPE with sPEPE. In particular, differences arose in the coordination of both Li^+ and $TFSI^-$ by the polymer when in the presence of the sPEPE polymer, as well as in the ion-pairing between the dopant species. Radial distributions and coordination plots were integral to portraying these structural changes. To this end, trajectories from the 2007 study were reanalyzed to present the data in the comparisons below.

Using the standard prescription[6] for calculating the radial distribution function and using the graphical visualization tool VMD[65], the plots in Figure 4.2 were constructed to show the interaction of Li^+ with a) $TFSI^-$, b) ether oxygens on the sidechain combs, c) backbone ether oxygens, as well as d) with itself. Closely related to the RDFs are the coordination numbers – an average measure of how many Li^+ cations surround other species in the simulation cell – which are presented in Table 4.2. Li^+ and $TFSI^-$ show a slightly increased propensity to be paired in the sPEPE system, demonstrated by the higher coordination seen in their RDF and a ~20% increase in first shell neighbors. Li^+ in the sPEPE system is also more strongly coordinated by the sidechain oxygens, while interacting drastically less (93% reduction in coordination) with the backbone oxygens compared with the PEPE SPE system. A reduced coordination with the backbone EOs leaves Li^+ more available to interact with $TFSI^-$ as the cation is

transported along the sidechains. This also explains the near absence of the first peak in the Li^+-Li^+ RDF for the sPEPE system, due to a 90% reduction in nearest Li^+-Li^+ neighbors. That peak is quite conspicuous in the PEPE system due to close Li^+ approach to opposing sides of individual backbone EOs.

The RDFs involving the N atom of $TFSI^-$ with a) other $TFSI^-$s, b) sidechain EOs, and c) backbone EOs are shown in Figure 4.3. $TFSI^-$ in the sPEPE system, similar to Li^+ albeit not as disparately, follows the trend of interacting more with the sidechains and less with the backbone compared to the PEPE system, simply due to "following" Li^+ to the cation's more preferential coordination location on the sidechains. Interestingly, a noticeable difference can be seen in the RDF of $TFSI^-$ with itself. The PEPE system allows TFSI anions to approach each other more closely, attributable to multiple $TFSI^-$s "sharing" the coordination with particular Li^+ cations along the course of the simulation trajectory. This phenomenon is absent in the sPEPE system.

Figure 4.4 shows RDFs relative to the carbon atoms on the sPEPE or PEPE chains and is useful for understanding the structural changes of the chain and its solvation when introducing the spacers. In the PEPE system, the single CH_2 linking the backbone and sidechains shows a strong association with Li^+ (Figure 4.4a) due to the latter's propensity to associate with the EOs close by. Li^+ is contrarily hardly associated with the first CH_2 group (C1) of the spacer in the sPEPE system, associating instead at the C4 end of the spacer near the first comb EO with a four-fold increase in first shell neighbors (Table 4.2). $TFSI^-$ shows little inclination for the spacer carbon atoms in either system (Figure 4.4b), yet shows a greater association with the C4 end of the sPEPE spacer. Li^+ shows a higher association with terminal CH_3 groups (Figure 4.4c) in the sPEPE system due to the

cation's higher availability away from the backbone, while $TFSI^-$ is also slightly more associated with these terminal groups in the sPEPE system. Finally, the morphology of the chain itself can be seen from the spacer carbon – spacer carbon RDFs in Figure 4.4d. The association of the C1 atoms of the spacers with others remains hardly perturbed by increasing the spacer length to four CH_2 groups in sPEPE, rotating through and maintaining mostly a trans orientation along the backbone, a fact consistent with an analysis of the backbone dihedrals (data not shown). Association of the C4 carbons across neighboring combs in the sPEPE SPE is minimal due to the high degree of entropy of the chain despite what would be a natural propensity of these aliphatic moieties to cluster together.

The coordination plot shown in Figure 4.5 is of particular use in understanding the differences in Li^+ interaction with the SPEs. For this analysis, the 8 ether oxygens in each backbone repeat segment were labeled 0 through 7 according to their position away from the backbone further down the sidechain. In effect, backbone oxygens were given a value of 0, while the furthest ether oxygen on a sidechain had a value of 7. At each timestep for each Li^+ along the trajectories, an average value of the assigned value of each coordinated oxygen, i.e. where $r(Li^+\text{-}EO) < 2.8$Å for that particular Li^+, was calculated. This allows an analysis to be seen of the "chain level" at which every Li^+ was coordinated over time, shown for several representative Li^+ cations in the left panels of Figure 4.5. Conclusions previously drawn from the PEPE system[41] showed Li^+ to exhibit a degraded transport being strongly coordinated to the backbone. This is confirmed by the coordination analysis seen in the top panels of Figure 4.5, showing a high probability for Li^+ to interact with lower levels of the polymer chain. In contrast, the sPEPE system

(bottom panel of Figure 4.5) shows that Li^+ cations spend little time coordinated with the backbone oxygens, and instead preferentially interact with the sidechain EOs.

Dynamical Differences Between Systems

Of main interest in this study is the effect of introducing the spacers to the SPE on the transport of Li^+ and its conductivity. With structural results showing the Li^+ to be significantly less coordinated by the backbone of the sPEPE SPE, Li^+ was expected to show an increased diffusion coefficient with respect to Li^+ of the PEPE system. Mean square displacements (MSDs) of the ionic species were calculated from the full trajectories, and are shown in Figure 4.6 for the temperatures 423, 393, and 363 K. These were used with the Einstein relation[6] to calculate the diffusion coefficient *D*. The ion self-diffusion coefficients are presented in Figure 4.7 and show Li^+ to diffuse faster and $TFSI^-$ to diffuse less rapidly for the sPEPE system across the temperature range studied compared to the PEPE results. With production run times approaching the order of 100 ns, $TFSI^-$ reached the diffusive regime across all temperatures, while Li^+ only did not at the lowest temperature studied, 363 K. To calculate D_{Li}(363 K), the value at 423 K was scaled by the temperature-dependent time-shift factor following the procedure outlined[41] for the low-temperature simulations performed on PEPE system by superimposing mean square displacements at those two temperatures.

In order to understand the reduced diffusion of $TFSI^-$ in the sPEPE system, it was beneficial to consider the dynamics of the SPE chains and their association with the TFSI anions. The addition of the $(CH_2)_4$ spacer makes the sPEPE chain bulkier and slower; analysis of the MSD of the sPEPE chain shows its speed to be reduced by about half during the course of the simulation than its PEPE counterpart. The extra bulk of the

lengthened spacers, in effect, creates a higher barrier for chain relaxation, resulting in stiffer backbones of the SPE chains. The effect of backbone stiffness has been probed with MD simulations. These studies show that artificially increasing the backbone stiffness has a slowing effect on ion transport due to the entire system being somewhat slower to relax.[41, 66] In particular for Li^+, when complexed by the artificially slowly relaxing backbone, the slow-down in ion transport resulted.[41] In the sPEPE system studied here, this limitation is overcome; Li^+ is negligibly coordinated to the sPEPE backbone and free to move about along the side chains, being unaffected by the more slowly relaxing backbone. A system-wide slowing, though not extreme, remains for the TFSI anion. The slower diffusion of TFSI anions compared to the PEPE system can be understood due to this and to their more probable association with the combs (Figure 3b) and carbon atoms of the spacer and CH_3 termini (Figure 4.4b and c).

To examine the effect of the sPEPE system on the conductivity, the same procedure was followed to calculate conductivity, λ, and the degree of uncorrelated ion motion, α_d, as reported previously.[41] These results are presented in Table 4.1. The ionic conductivites of the PEPE systems are slightly higher (~10%) across the temperature ranges studied compared to the sPEPE systems, due to the decreased diffusion of the TFSI anions. The conductivity of the Li^+ cations, however, is two-to-three times higher for sPEPE vs. that of the PEPE systems. The trends seen in the PEPE system for an increasing fraction of free ions and increasing degree of ion-uncorrelated motion with decreasing temperature are also seen for the sPEPE system. As seen in the Li^+-$TFSI^-$ radial distribution function (Figure 4.2a) and coordination numbers (Table 4.2), evidence

for Li^+ and $TFSI^-$ to be more paired in the sPEPE system is also witnessed in the higher values of fraction of free ions.

The average time that Li^+ cations and TFSI anions stay associated, the residence time, is also insightful as to their behavior in the system. Using the methodology of Laage and Hynes,[67] the probability of dissociating LiTFSI was calculated as a function of time t and fit to the exponential $\exp(-t/\tau)$. Integrating this function to infinity produced the values for the residence time found in Table 4.3. Individual LiTFSIs remained associated longer in the PEPE system at comparable temperatures for sPEPE. Although greater pairing was present in the sPEPE system, these pairs were switching more rapidly due to the more mobile nature of Li^+ cations, of benefit for the conductivity of the system.

Conclusions

The PEPE comb-branched polymer electrolyte was previously shown[41] to have a degraded ability to transport Li^+ compared to similar PEO electrolyte systems. In the present study, a novel modification was made to the PEPE polymer by incorporating non-coordinating CH_2 tethers at the attachment points between the backbone and the sidechains. This was done in an attempt to decouple the movement of the Li^+ along the sidechains, the desirable paradigm of ideal conduction of charge in the system, from long, unfavorable association with the slow-moving backbone. As seen in Figure 4.1, when the SPE PEPE is used (n=1, vs. n=4 for sPEPE), the oxygen-rich branch points can be an ideal location for Li^+ to be coordinated, which can be seen in a high amount of coordination at the lower chain level (top panel of Figure 4.5) and results in a degraded value for the diffusion of Li^+.

Using the sPEPE SPE, the main limitation of the PEPE system is overcome; increasing the distance between the backbone and sidechains with non-coordinating spacers results in very little coordination with the backbone, as seen in the bottom panel of Figure 4.5. This is confirmed further by the radial distribution functions between Li^+ and the backbone oxygens (Figure 4.2c). The result is a system that allows Li^+ to diffuse three times more quickly at the highest temperature studied than the original PEPE system. Although the ionic conductivity of the sPEPE system is slightly lower due to slower moving TFSI anions, the conductivity of the Li^+ is significantly improved and promising for future studies with this PEO-based comb-branched architecture.

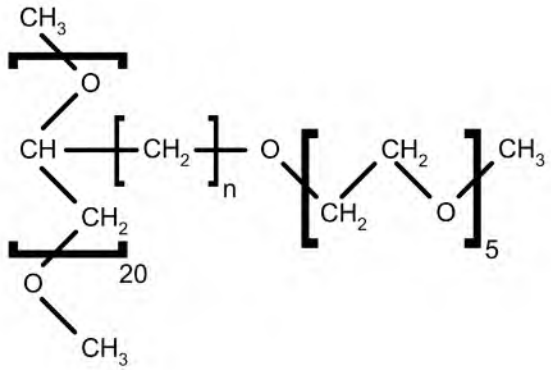


Figure 4.1. Chemical structure of the comb-branched polymer used in simulations for the present study of sPEPE (n=4) and PEPE (n=1).

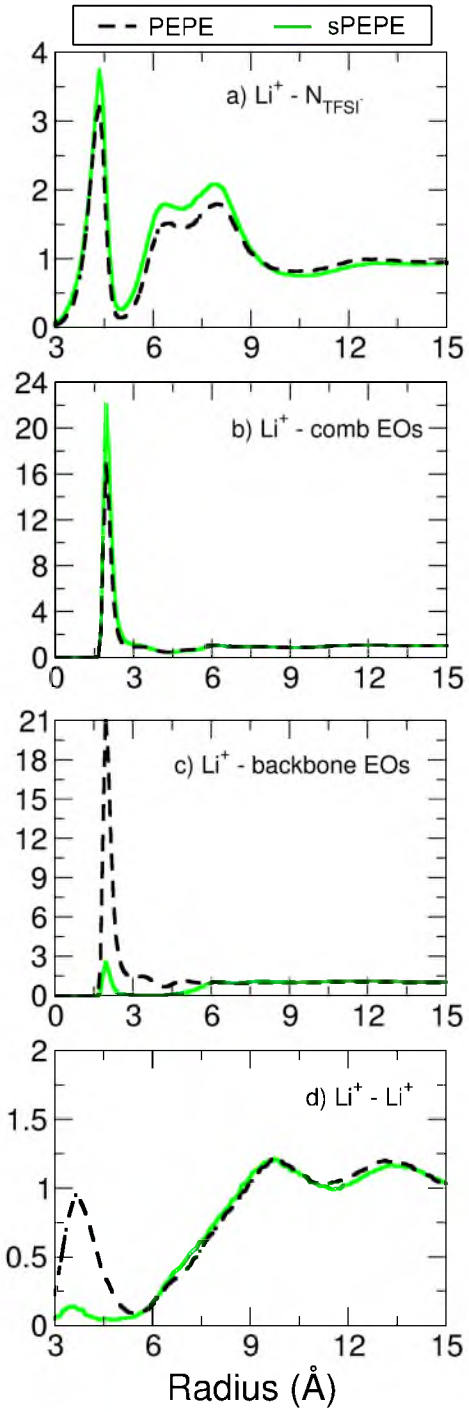


Figure 4.2. Radial distribution functions of Li^+ with a) the N atom of $TFSI^-$; b) ether oxygens on the combs of the polymer electrolyte; c) ether oxygens on the backbone of the polymer electrolyte; and d) with other Li^+ ions. sPEPE system data shown in solid green, PEPE in dashed black lines.

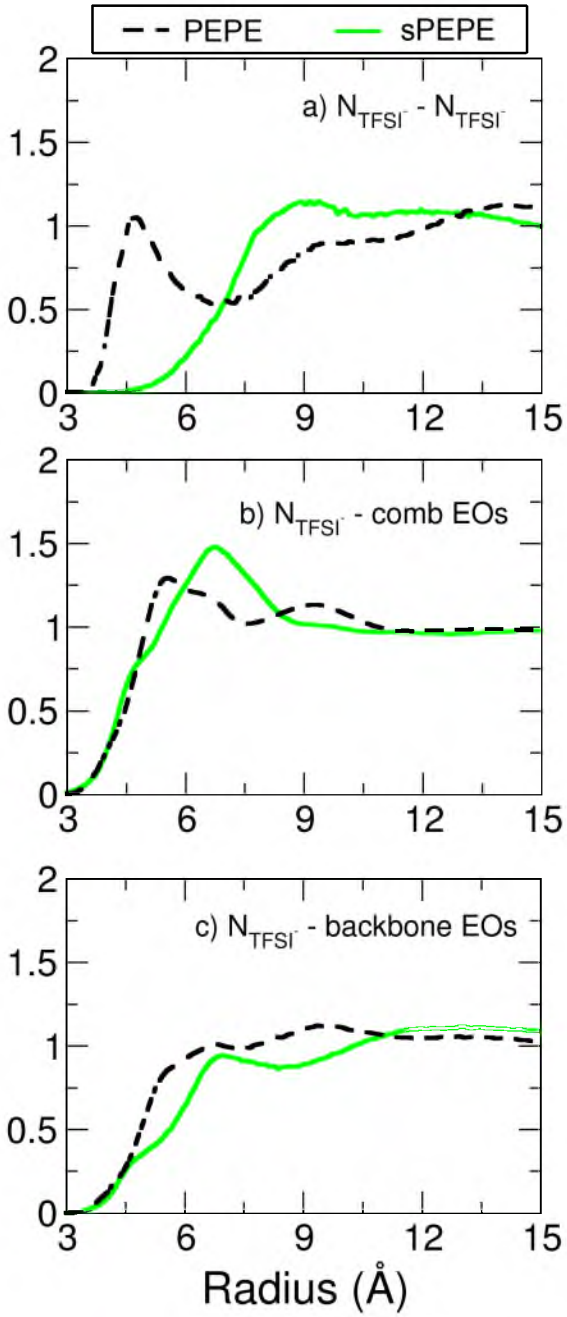


Figure 4.3. Radial distribution functions of the N atom of $TFSI^-$ with a) the N atom of other TFSI anions; b) ether oxygens on the combs of the polymer electrolyte; and c) ether oxygens on the backbone of the polymer electrolyte. sPEPE system data shown in solid green, PEPE in dashed black lines.

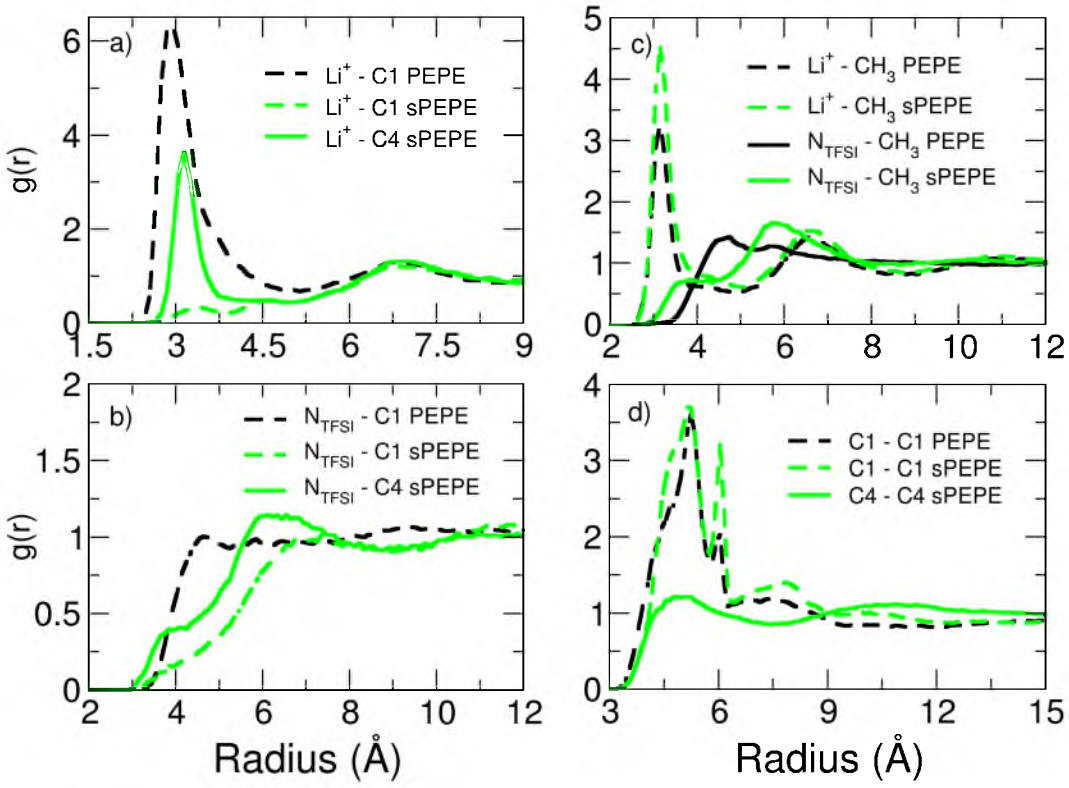


Figure 4.4. Radial distribution functions of a) the Li^+ cations and b) N atom of the TFSI anions with the C atoms of the backbone-to-sidechain connecting spacers of the PEPE and sPEPE SPE chains and c) with the C atom of terminal methyl groups on the chains. C1 refers to the carbon atom of the first (and only in the case of PEPE since n=1) CH_2 in Figure 4.1 above, while C4 is for the n=4 case for sPEPE. Panel d) shows the radial distribution functions of the C1 and C4 atoms of the spacers with other C1 and C4 atoms on different intra- or intermolecular combs.

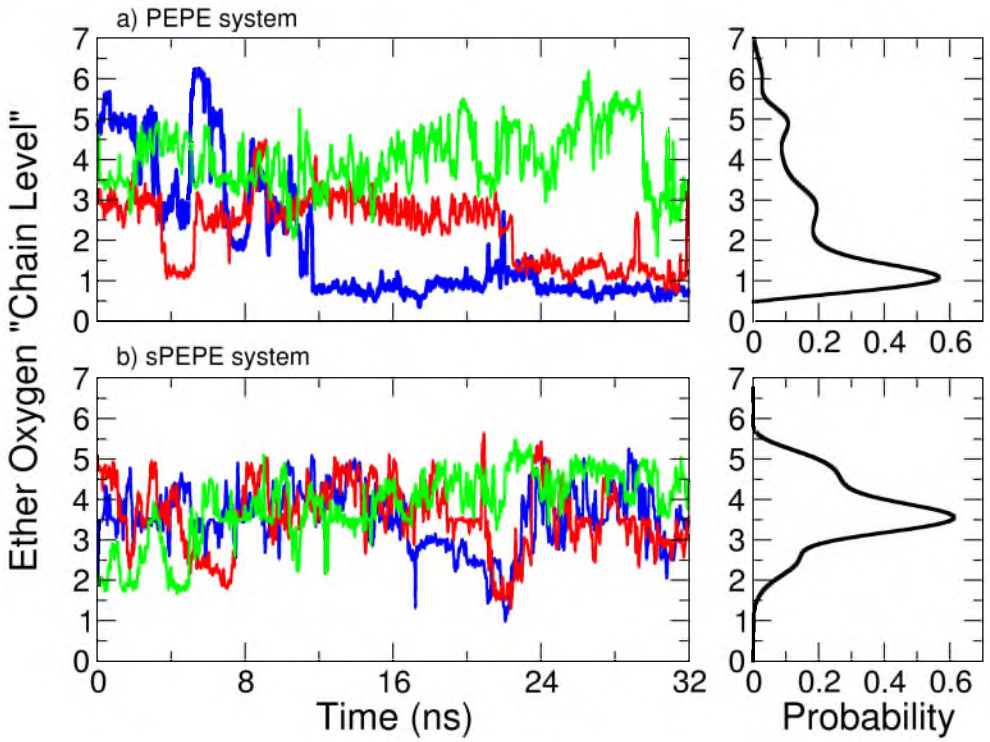


Figure 4.5. Simulation results at 423K of coordination of Li^+ cations with comb-branched polymer ether oxygens, which have been numbered sequentially from "chain level 0" (backbone EOs) to "chain level 7" (last EO on a comb before its terminus). Left panels show evolution of coordination over time for representative Li^+ cations for (a) PEPE systems and (b) the present study of sPEPE. Right panels show the probability distribution for coordination of all Li^+ cations with the EOs for the respective study.

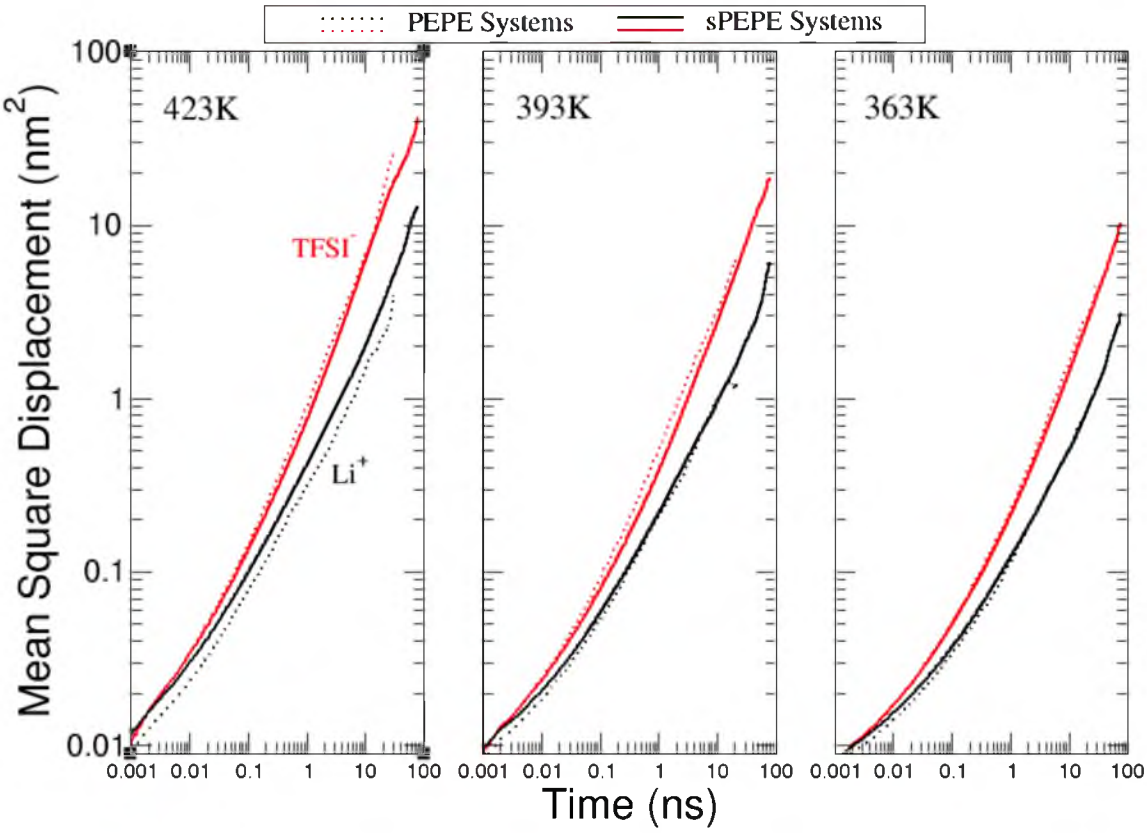


Figure 4.6. Average mean-square displacements calculated for simulations of the comb-branched polymer electrolyte systems for the TFSI anions (red) and Li^+ cations (black) at varying temperatures for sPEPE (solid lines) and PEPE systems (dotted lines).

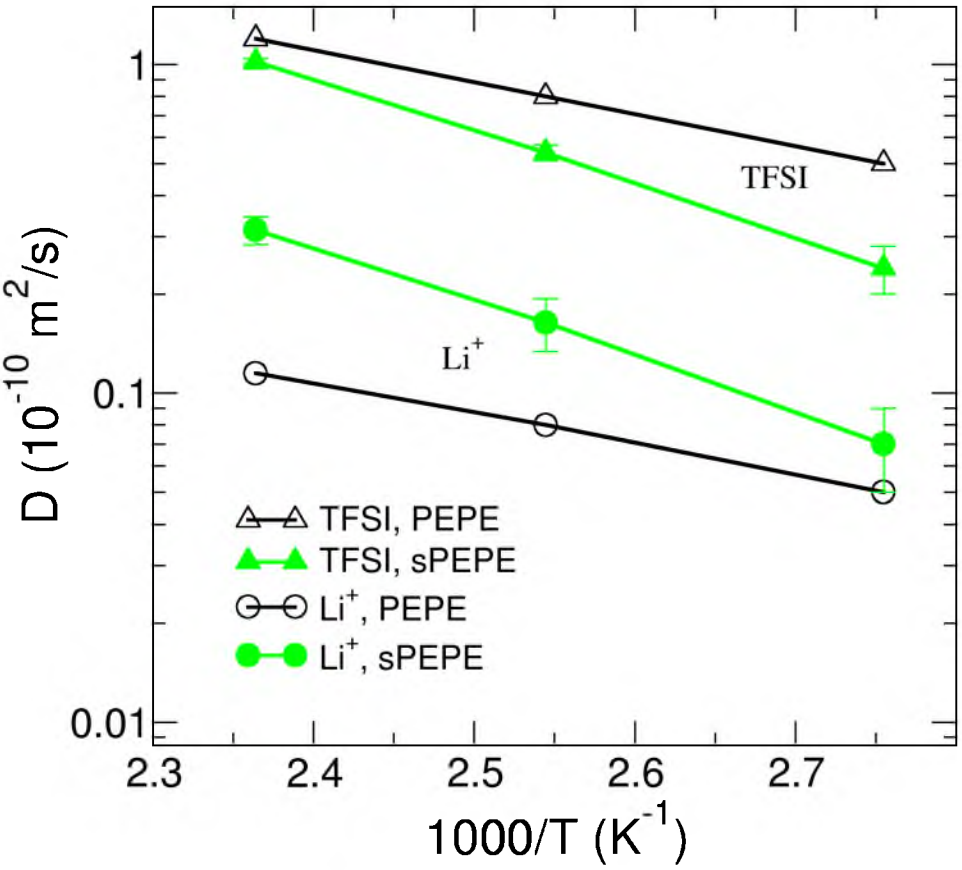


Figure 4.7. Li^+ (circle) and $TFSI^-$ (triangle) self-diffusion coefficients for the comb-branched polymer electrolyte sPEPE (black) and PEPE (green) systems.

Table 4.1. Simulation details and calculated conductivity data for PEPE and sPEPE systems

Temp (K)	*Simulation time (ns)*	*Ionic Conductivity λ (mS/cm)*	λ_{Li+} *(mS/cm)*	*Fraction of free ions*	*Degree of Ion Uncorrelated Motion*
PEPE system[a]					
423	38	3.3	0.29	0.80	0.83
393	24	1.8	0.22	0.88	0.86
363	30	0.9	0.04	0.89	0.90
sPEPE system					
423	80	3.0	0.70	0.83	0.81
393	80	1.7	0.41	0.84	0.80
363	80	0.8	0.19	0.85	0.97

a) Data recalculated from Ref. [41].

Table 4.2. Radial distribution g(r) data for the PEPE and sPEPE simulations at 423 K.

	r (Å) of first shell peak	*g(r) value at peak*	*r (Å) of extent of first shell*	*First shell neighbors*
Li^+-N_{TFSI-}				
PEPE	4.3	3.3	4.9	0.24
sPEPE	4.3	3.7	4.9	0.29
Li^+-Backbone Os				
PEPE	1.9	21.2	4.2	1.3
sPEPE	1.9	2.6	4.4	0.09
Li^+-Sidechain Os				
PEPE	1.9	16.8	4.3	5.4
sPEPE	1.9	22.1	4.3	5.8
Li^+- Li^+				
PEPE	3.6	0.95	5.4	0.13
sPEPE	3.5	0.14	4.9	0.014
Li^+-C_{Spacer}				
PEPE C1	3.1	4.9	5.0	0.47
sPEPE C1	3.4	0.33	5.0	0.12
sPEPE C4	3.1	3.6	5.0	0.60

Table 4.3. Residence time (ns) of LiTFSI pairing for the PEPE and sPEPE systems

Temp (K)	*PEPE*	*sPEPE*
423	1.58	0.80
393	1.96	1.03
363	5.83	3.24

REFERENCES

1. Levitt, M.; Warshel, A. *Nature* **1975,** 253, (5494), 694-698.

2. Plimpton, S. *Journal of Computational Physics* **1995,** 117, (1), 1-19.

3. Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *Journal of Computational Chemistry* **2005,** 26, (16), 1668.

4. Ayyagari, C.; Bedrov, D.; Borodin, O.; Smith, G. D. *Lucretius*, MD simulation code http://lucretius.mse.utah.edu.

5. Norberg, J.; Nilsson, L. *Biophysical Journal* **2000,** 79, (3), 1537-1553.

6. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*. 1st ed.; Clarendon Press: Oxford, 1987.

7. Ewald, P. *Annalen der Physik (Leipzig)* **1921,** 64, 253.

8. Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993,** 98, (12), 4.

9. Hockney, R. W.; Eastwood, J. W., *Computer Simulation Using Particles*. IOP, Bristol: 1988.

10. Adams, D. J.; McDonald, I. R. *Molecular Physics* **1976,** 32, (4), 931-947.

11. de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **1980,** 373, (1752), 27-56.

12. De Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Annual Review of Physical Chemistry* **1986,** 37, (1), 245-270.

13. Perram, J. W.; Petersen, H. G.; De Leeuw, S. W. *Molecular Physics* **1988,** 65, (4), 875 - 893.

14. Carre, A.; Berthier, L.; Horbach, J.; Ispas, S.; Kob, W. *Journal of Chemical Physics* **2007,** 127.

15. Muller-Plathe, F. *Chemphyschem* **2002,** 3, (9), 754-769.

16. Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Muller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V., Bridging the gap between atomistic and coarse-grained models of polymers: Status and perspectives. In *Advances in Polymer Science: Viscoelasticity, Atomistic Models, Statistical Chemistry*, Springer-Verlag Berlin: Berlin, 2000; Vol. 152, pp 41-156.

17. Molinero, V.; Goddard, W. A. *Journal of Physical Chemistry B* **2004,** 108, (4), 1414-1427.

18. Noid, W. G.; Chu, J. W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. *Journal of Chemical Physics* **2008,** 128, (24), 11.

19. Noid, W. G.; Liu, P.; Wang, Y.; Chu, J. W.; Ayton, G. S.; Izvekov, S.; Andersen, H. C.; Voth, G. A. *Journal of Chemical Physics* **2008,** 128, (24), 20.

20. Voth, G. A., *Coarse-Graining of Condensed Phase and Biomolecular Systems*. Crc Press-Taylor & Francis Group: 2009; p 1-4.

21. Bennun, S. V.; Hoopes, M. I.; Xing, C. Y.; Faller, R. *Chemistry and Physics of Lipids* **2009,** 159, (2), 59-66.

22. Marrink, S. J.; de Vries, A. H.; Tieleman, D. P. *Biochimica Et Biophysica Acta-Biomembranes* **2009,** 1788, (1), 149-168.

23. Jorgensen, W. L.; Tirado-Rives, J. *Journal of the American Chemical Society* **1988,** 110, (6), 1657-1666.

24. Molinero, V.; Moore, E. *Journal of Physical Chemistry B* **2009,** 113, (13), 4008-4016.

25. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *Journal of Chemical Physics* **2004,** 120, (20), 9665-9678.

26. Stillinger, F. H.; Weber, T. A. *Physical Review B* **1985,** 31, (8), 5262.

27. Jacobson, L. C.; Hujo, W.; Molinero, V. *Journal of the American Chemical Society* **2010,** 132, (33), 11806.

28. de la Llave, E.; Molinero, V.; Scherlis, D. A. *The Journal of Physical Chemistry C* **2011,** 116, (2), 1833-1840.

29. González Solveyra, E.; de la Llave, E.; Scherlis, D. A.; Molinero, V. *The Journal of Physical Chemistry B* **2011,** 115, (48), 14196-14204.

30. Johnston, J. C.; Molinero, V. *Journal of the American Chemical Society* **2012,** 134, (15), 6650-6659.

31. Limmer, D. T.; Chandler, D. *The Journal of Chemical Physics* **2011,** 135, (13), 134503-10.

32. Cadena, C.; Anthony, J. L.; Shah, J. K.; Morrow, T. I.; Brennecke, J. F.; Maginn, E. J. *Journal of the American Chemical Society* **2004,** 126, (16), 5300-5308.

33. Smith, G. D.; Borodin, O.; Li, L.; Kim, H.; Liu, Q.; Bara, J. E.; Gin, D. L.; Nobel, R. *Physical Chemistry Chemical Physics* **2008,** 10, (41), 6301-6312.

34. Cadena, C.; Zhao, Q.; Snurr, R. Q.; Maginn, E. J. *The Journal of Physical Chemistry B* **2006,** 110, (6), 2821-2832.

35. Cadena, C.; Maginn, E. J. *The Journal of Physical Chemistry B* **2006,** 110, (36), 18026-18039.

36. Borodin, O. *The Journal of Physical Chemistry. B* **2009,** 113, (33).

37. Bedrov, D.; Borodin, O.; Li, Z.; Smith, G. D. *The Journal of Physical Chemistry. B* **2010,** 114, (15).

38. Bhargava, B. L.; Balasubramanian, S. *The Journal of Chemical Physics* **2007,** 127, (11), 114510-6.

39. Micaelo, N. M.; Baptista, A. M.; Soares, C. M. *The Journal of Physical Chemistry B* **2006,** 110, (29), 14444-14451.

40. Borodin, O.; Smith, G. D. *Macromolecules* **2006,** 39, (4), 1620-1629.

41. Borodin, O.; Smith, G. D. *Macromolecules* **2007,** 40, (4), 1252-1258.

42. Fergus, J. W. *Journal of Power Sources* **2010,** 195, (15), 4554-4569.

43. Quartarone, E.; Mustarelli, P. *Chemical Society Reviews* **2011,** 40, (5), 2525-2540.

44. Palacin, M. R. *Chemical Society Reviews* **2009,** 38, (9), 2565-2575.

45. Jacob, M. M. E.; Hackett, E.; Giannelis, E. P. *Journal of Materials Chemistry* **2003,** 13, (1), 1-5.

46. Liang, W.-J.; Chen, Y.-P.; Wu, C.-P.; Kuo, P.-L. *The Journal of Physical Chemistry. B* **2005,** 109, (51), 24311-8.

47. Armand, M.; Tarascon, J. M. *Nature* **2008,** 451, (7179), 652-7.

48. Dunn, B.; Kamath, H.; Tarascon, J.-M. *Science* **2011,** 334, (6058), 928-935.

49. Arico, A. S.; Bruce, P.; Scrosati, B.; Tarascon, J.-M.; van Schalkwijk, W. *Nature Materials* **2005,** 4, (5), 366-377.

50. Buriez, O.; Han, Y. B.; Hou, J.; Kerr, J. B.; Qiao, J.; Sloop, S. E.; Tian, M.; Wang, S. *Journal of Power Sources* **2000,** 89, (2), 149-155.

51. Kerr, J. B.; Sloop, S. E.; Liu, G.; Han, Y. B.; Hou, J.; Wang, S. *Journal of Power Sources* **2002,** 110, (2), 389-400.

52. Liu, G.; Reeder, C. L.; Sun, X.; Kerr, J. B. *Solid State Ionics* **2004,** 175, (1-4), 781-783.

53. Sun, X. G.; Hou, J.; Kerr, J. B. *Electrochimica Acta* **2005,** 50, (5), 1139-1147.

54. Sun, X. G.; Kerr, J. B. *Macromolecules* **2006,** 39, (1), 362-372.

55. Munch Elmér, A.; Jannasch, P. *Solid State Ionics* **2006,** 177, (5–6), 573-579.

56. Kaskhedikar, N.; Paulsdorf, J.; Burjanadze, M.; Karatas, Y.; Wilmer, D.; Roling, B.; Wiemhöfer, H. D. *Solid State Ionics* **2006,** 177, (7–8), 703-707.

57. Luther, T. A.; Stewart, F. F.; Budzien, J. L.; LaViolette, R. A.; Bauer, W. F.; Harrup, M. K.; Allen, C. W.; Elayan, A. *The Journal of Physical Chemistry B* **2003,** 107, (14), 3168-3176.

58. Allcock, H. R.; Napierala, M. E.; Olmeijer, D. L.; Best, S. A.; Merz, K. M. *Macromolecules* **1999,** 32, (3), 732-741.

59. Watanabe, M.; Hirakimoto, T.; Mutoh, S.; Nishimoto, A. *Solid State Ionics* **2002,** 148, (3–4), 399-404.

60. Balbuena, P. B.; Lamas, E. J.; Wang, Y. *Electrochimica Acta* **2005,** 50, (19), 3788-3795.

61. Borodin, O.; Smith, G. D. *The Journal of Physical Chemistry. B* **2006,** 110, (12), 6293-9.

62. Borodin, O.; Smith, G. D. *The Journal of Physical Chemistry. B* **2006,** 110, (12), 6279-92.

63. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Molecular Physics* **1996,** 87, (5), 1117-1157.

64. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Computational Physics* **1977,** 23, (3), 327-341.

65. Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996,** 14, (1), 33-38.

66. Borodin, O.; Smith, G. D. *The Journal of Physical Chemistry. B* **2006,** 110, (10), 4971-7.

67. Laage, D.; Hynes, J. T. *The Journal of Physical Chemistry. B* **2008,** 112, (26), 7697-7701.